SPECTRAL VARIATIONAL INTEGRATORS

JAMES HALL AND MELVIN LEOK

ABSTRACT. In this paper, we present a new variational integrator for problems in Lagrangian mechanics. Using techniques from Galerkin variational integrators, we construct a scheme for numerical integration that converges geometrically, and is symplectic and momentum preserving. Furthermore, we prove that under appropriate assumptions, variational integrators constructed using Galerkin techniques will yield numerical methods that are in a certain sense optimal, converging at the same rate as the best possible approximation in a certain function space. We further prove that certain geometric invariants also converge at an optimal rate, and that the error associated with these geometric invariants is independent of the number of steps taken. We close with several numerical examples that demonstrate the predicted rates of convergence.

1. INTRODUCTION

There has been significant recent interest in the development of structure-preserving numerical methods for variational problems. One of the key points of interest is developing high-order symplectic integrators for Lagrangian systems. The generalized Galerkin framework has proven to be a powerful theoretical and practical tool for developing such methods. This paper presents a high-order Galerkin variational integrator that exhibits geometric convergence to the true flow of a Lagrangian system. In addition, this method is symplectic, momentum-preserving, and stable even for very large time steps.

Galerkin variational integrators fall into the general framework of discrete mechanics. For a general and comprehensive introduction to the subject, the reader is referred to Marsden and West [19]. Discrete mechanics develops mechanics from discrete variational principles, and, as Marsden and West demonstrated, gives rise to many discrete structures which are analogous to structures found in classical mechanics. By taking these structures into account, discrete mechanics suggests numerical methods which often exhibit excellent long term stability and qualitative behavior. Because of these qualities, much recent work has been done on developing numerical methods from the discrete mechanics viewpoint. See, for example, Hairer et al. [9] for a broad overview of the field of geometric numerical integration, and Marsden and West [19]; Müller and Ortiz [22]; Patrick and Cuell [24] discuss the error analysis of variational integrators. Various extensions have also been considered, including, Lall and West [10]; Leok and Zhang [17] for Hamiltonian systems; Fetecau et al. [8] for nonsmooth problems with collisions; Lew et al. [18]; Marsden et al. [20] for Lagrangian PDEs; Cortés and Martínez [5]; Fedorov and Zenkov [7]; McLachlan and Perlmutter [21] for nonholonomic systems; Bou-Rabee and Owhadi [2, 3] for stochastic Hamiltonian systems; Bou-Rabee and Marsden [1]; Lee et al. [12, 13] for problems on Lie groups and homogeneous spaces.

The fundamental object in discrete mechanics is the discrete Lagrangian $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$, where Q is a configuration manifold. The discrete Lagrangian is chosen to be an approximation to the action of a Lagrangian over the time step $[0, h]$,

$$L_d(q_0, q_1, h) \approx \underset{\substack{q \in C^2([0,h], Q) \\ q(0)=q_0, q(h)=q_1}}{\text{ext}} \int_0^h L(q, \dot{q}) \, dt,$$

or simply $L_d(q_0, q_1)$ when h is assumed to be constant. Discrete mechanics is formulated by finding stationary points of a discrete action sum based on the sum of discrete Lagrangians,

$$\mathbb{S}(\{q_k\}_{k=1}^n) = \sum_{k=1}^{n-1} L_d(q_k, q_{k+1}) \approx \int_{t_1}^{t_2} L(q, \dot{q}) \, dt.$$

For Galerkin variational integrators specifically, the discrete Lagrangian is induced by constructing a discrete approximation of the action integral over the interval $[0, h]$ based on a finite-dimensional function space and quadrature rule. Once this discrete action is constructed, the discrete Lagrangian can be recovered by solving

for stationary points of the discrete action subject to fixed endpoints, and then evaluating the discrete action at these stationary points,

$$
(1) \qquad L_d\left(q_0, q_1, h\right) = \operatorname*{ext}_{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_1}} h \sum_{j=1}^{m} b_j L\left(q\left(c_j h\right), \dot{q}\left(c_j h\right)\right).
$$

Because the rate of convergence of the approximate flow to the true flow is related to how well the discrete Lagrangian approximates the true action, this type of construction gives a method for constructing and analyzing high-order methods. The hope is that the discrete Lagrangian inherits the accuracy of the function space used to construct it, much in the same way as standard finite-element methods. We will show that for certain Lagrangians, Galerkin constructions based on high-order approximation spaces do in fact result in correspondingly high order methods.

Significant work has already been done constructing and analyzing these types of Galerkin variational integrators. In Leok [14], a number of different possible constructions based on the Galerkin framework are presented. In Leok and Shingel [15], Hermite polynomials are used to construct globally smooth high-order methods. What separates this work from the work that precedes it is the use of a spectral approximation paradigm, which induces methods that exhibit geometric convergence. This type of convergence is established theoretically and demonstrated through numerical examples.

1.1. **Discrete Mechanics.** Before discussing the construction and convergence of spectral variational integrators, it is useful to review some of the fundamental results from discrete mechanics that are used in our analysis. We have already introduced the *discrete Lagrangian* $L_d : Q \times Q \times \mathbb{R} \to \mathbb{R}$,

$$
L_d\left(q_0, q_1, h\right) \approx \operatorname*{ext}_{\substack{q \in C^2([0,h],Q) \\ q(0)=q_0, q(h)=q_1}} \int_0^h L\left(q, \dot{q}\right) \mathrm{d}t.
$$

and the *discrete action sum*,

$$
\mathbb{S}\left(\{q_k\}_{k=1}^n\right) = \sum_{k=1}^{n-1} L_d\left(q_k, q_{k+1}\right) \approx \int_{t_1}^{t_2} L\left(q, \dot{q}\right) \mathrm{d}t.
$$

Taking variations of the discrete action sum and using discrete integration by parts leads to the discrete Euler-Lagrange equations,

$$
(2) \qquad D_2 L_d\left(q_{k-1}, q_k\right) + D_1 L_d\left(q_k, q_{k+1}\right) = 0,
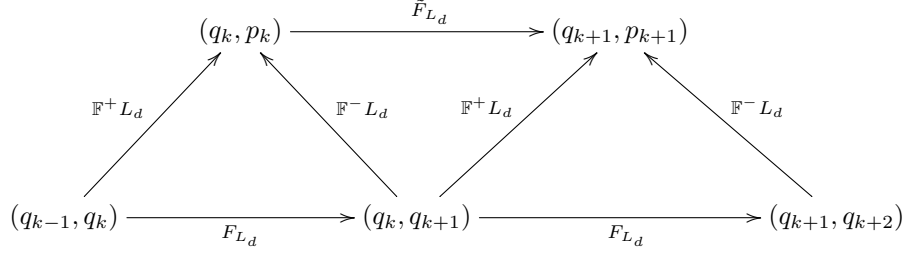$$

where D_1 denotes differentiation with respect to the first argument and D_2 denotes differentiation with respect to the second argument. Given (q_{k-1}, q_k), these equations implicitly define an update map, known as the *discrete Lagrangian flow map*, $F_{L_d} : Q \times Q \to Q \times Q$, given by $F_{L_d}\left(q_{k-1}, q_k\right) = \left(q_k, q_{k+1}\right)$, where $(q_{k-1}, q_k), (q_k, q_{k+1})$ satisfy (2). Furthermore, the discrete Lagrangian defines the *discrete Legendre transforms*, $\mathbb{F}^{\pm} L_d : Q \times Q \to T^*Q$:

$$
\mathbb{F}^+ L_d : \left(q_0, q_1\right) \to \left(q_1, p_1\right) = \left(q_1, D_2 L_d\left(q_0, q_1\right)\right),
$$
$$
\mathbb{F}^- L_d : \left(q_0, q_1\right) \to \left(q_0, p_0\right) = \left(q_0, -D_1 L_d\left(q_0, q_1\right)\right).
$$

Using the discrete Legendre transforms, we define the *discrete Hamiltonian flow map*, $\tilde{F}_{L_d} : T^*Q \to T^*Q$,

$$
\tilde{F}_{L_d} : \left(q_0, p_0\right) \to \left(q_1, p_1\right) = \mathbb{F}^+ L_d\left(\left(\mathbb{F}^- L_d\right)^{-1}\left(q_0, p_0\right)\right).
$$

2

The following commutative diagram illustrates the relationship between the discrete Hamiltonian flow map, discrete Lagrangian flow map, and the discrete Legendre transforms,

$$
\begin{array}{ccccc}
(q_k, p_k) & \xrightarrow{\tilde{F}_{L_d}} & (q_{k+1}, p_{k+1}) \\
\mathbb{F}^+ L_d \nearrow \quad \nwarrow \mathbb{F}^- L_d & \quad \mathbb{F}^+ L_d \nearrow \quad \nwarrow \mathbb{F}^- L_d \\
(q_{k-1}, q_k) \xrightarrow{F_{L_d}} (q_k, q_{k+1}) \xrightarrow{F_{L_d}} (q_{k+1}, q_{k+2})
\end{array}
$$

We now introduce the *exact discrete Lagrangian* L_d^E,

$$
L_d^E (q_0, q_1, h) = \underset{\substack{q(0)=q_k \\ q(h)=q_{k+1} \quad q_0 q_1 \\ q \in C^2([0,h], Q)}}{\text{ext}} \int_0^h L(q, \dot{q}) \, dt.
$$

An important theoretical result for the error analysis of variational integrators is that the discrete Hamiltonian and Lagrangian flow maps associated with the exact discrete Lagrangian produces an exact sampling of the true flow, as was shown in Marsden and West [19]. Using this result, Marsden and West [19] shows that there is a fundamental relationship between how well a discrete Lagrangian L_d approximates the exact discrete Lagrangian L_d^E and how well the corresponding discrete Hamiltonian flow maps, discrete Lagrangian flow maps and discrete Legendre transforms approximate each other. Since the exact discrete Lagrangian produces an exact sampling of the true flow, this in turn leads to the following theorem regarding the error analysis of variational integrators, also found in Marsden and West [19]:

Theorem 1.1. (Variational Error Analysis) *Given a regular Lagrangian L and corresponding Hamiltonian H, the following are equivalent for a discrete Lagrangian L_d:*

(1) *the discrete Hamiltonian flow map for L_d has error $\mathcal{O}\left(h^{p+1}\right)$,*
(2) *the discrete Legendre transforms of L_d have error $\mathcal{O}\left(h^{p+1}\right)$,*
(3) *L_d approximates the exact discrete Lagrangian with error $\mathcal{O}\left(h^{p+1}\right)$.*

We will make extensive use of this theorem later when we analyze the convergence of spectral variational integrators.

In addition, in Marsden and West [19], it is shown that integrators constructed in this way, which are referred to as *variational integrators*, have significant geometric structure. Most importantly, variational integrators always conserve the canonical symplectic form, and a discrete Noether's Theorem guarantees that a discrete momentum map is conserved for any continuous symmetry of the discrete Lagrangian. The preservation of these discrete geometric structures underlie the excellent long term behavior of variational integrators.

2. CONSTRUCTION

2.1. **Generalized Galerkin Variational Integrators.** The construction of spectral variational integrators falls within the framework of generalized Galerkin variational integrators, discussed in Leok [14] and Marsden and West [19]. The motivating idea is that we try to replace the generally non-computable exact discrete Lagrangian $L_d^E (q_k, q_{k+1})$ with a computable discrete analogue, $L_d^G (q_k, q_{k+1})$. Galerkin variational integrators are constructed by using a finite-dimensional function space to discretize the action of a Lagrangian. Specifically, given a Lagrangian $L : TQ \to \mathbb{R}$, to construct a Galerkin variational integrator:

(1) choose an n-dimensional function space $\mathbb{M}^n([0,h], Q) \subset C^2([0,h], Q)$, with a finite set of basis functions $\{\phi_i(t)\}_{i=1}^n$,
(2) choose a quadrature rule $\mathcal{G}(\cdot) : F([0,h], \mathbb{R}) \to \mathbb{R}$, so that $\mathcal{G}(f) = h \sum_{j=1}^m b_j f(c_j h) \approx \int_0^h f(t) \, dt$, where F is some appropriate function space,
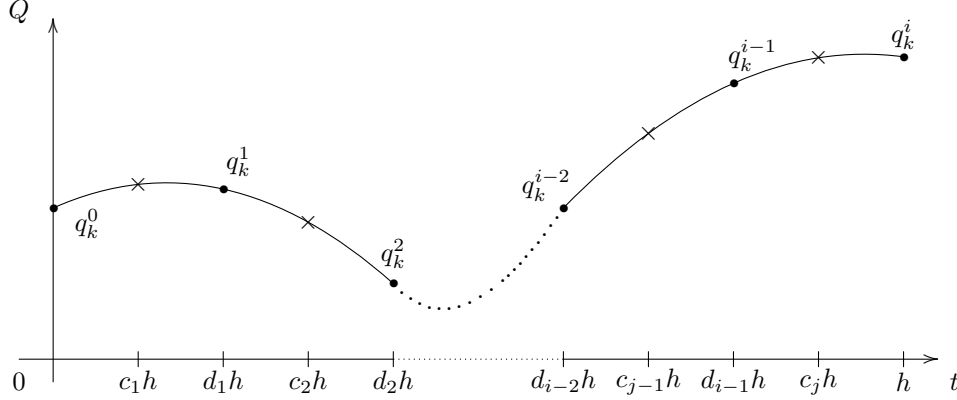
FIGURE 1. A visual schematic of the curve $\tilde{q}_n(t) \in \mathbb{M}^n([0, h], Q)$. The points marked with ($\times$) represent the quadrature points, which may or may not be the same as interpolation points $d_i h$. In this figure we have chosen to depict a curve constructed from interpolating basis functions, but this is not necessary in general.

and then construct the discrete action $\mathbb{S}_d \left(\{q_k^i\}_{i=1}^n \right) : \prod_{i=1}^n Q_i \to \mathbb{R}$, (not to be confused with the discrete action sum $\mathbb{S} \left(\{q_k\}_{k=1}^\infty \right)$),

$$\mathbb{S}_d \left(\{q_k^i\}_{i=1}^n \right) = \mathcal{G} \left(L \left(\sum_{i=1}^n q_k^i \phi_i(t), \sum_{i=1}^n q_k^i \dot{\phi}_i(t) \right) \right) = h \sum_{j=1}^m b_j L \left(\sum_{i=1}^n q_k^i \phi_i(c_j h), \sum_{i=1}^n q_k^i \dot{\phi}_i(c_j h) \right),$$

where we use superscripts to index the weights associated with each basis function, as in Marsden and West [19].

Once the discrete action has been constructed, a discrete Lagrangian can be induced by finding stationary points $\tilde{q}_n(t) = \sum_{i=1}^n q_k^i \phi_i(t)$ of the action under the conditions $\tilde{q}_n(0) = \sum_{i=1}^n q_k^i \phi_i(0) = q_k$ and $\tilde{q}_n(h) = \sum_{i=1}^n q_k^i \phi_i(h) = q_{k+1}$ for some given q_k and q_{k+1},

$$L_d(q_k, q_{k+1}, h) = \underset{\substack{\tilde{q}_n(0)=q_k \\ \tilde{q}_n(h)=q_{k+1}}}{\text{ext}} \mathbb{S}_d \left(\{q_k^i\}_{i=1}^n \right) = \underset{\substack{\tilde{q}_n(0)=q_k \\ \tilde{q}_n(h)=q_{k+1}}}{\text{ext}} h \sum_{j=1}^m b_j L \left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h) \right).$$

A discrete Lagrangian flow map that result from this type of discrete Lagrangian is referred to as a Galerkin variational integrator.

2.2. **Spectral Variational Integrators.** There are two defining features of spectral variational integrators. The first is the choice of function space $\mathbb{M}^n([0, h], Q)$, and the second is that convergence is achieved not by shortening the time step h, but by increasing the dimension n of the function space.

2.2.1. *Choice of Function Space.* Restricting our attention to the case where Q is a linear space, spectral variational integrators are constructed using the basis functions $\phi_i(t) = l_i(t)$, where $l_i(t)$ are Lagrange interpolating polynomials based on the points $h_i = \frac{h}{2} \cos\left(\frac{i\pi}{n}\right) + \frac{h}{2}$ which are the Chebyshev points $t_i = \cos\left(\frac{i\pi}{n}\right)$, rescaled and shifted from $[-1, 1]$ to $[0, h]$. The resulting finite dimensional function space $\mathbb{M}^n([0, h], Q)$ is simply the polynomials of degree at most n on Q. However, the choice of this particular set of basis functions offer several advantages over other possible bases for the polynomials:

(1) the restriction on variations $\sum_{i=1}^n \delta q_k^i \phi_i(0) = \sum_{i=1}^n \delta q_k^i \phi_i(h) = 0$ reduces to $\delta q_k^1 = \delta q_k^n = 0$,
(2) the condition $\tilde{q}_n(0) = q_k$ reduces to $q_k^1 = q_k$,
(3) the induced numerical methods have generally better stability properties because of the excellent approximation properties of the interpolation polynomials at the Chebyshev points.

4

Using this choice of basis functions, for any chosen quadrature rule, the discrete Lagrangian becomes,

$$L_d(q_k, q_{k+1}, h) = \underset{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_k^1 = q_k, q_k^n = q_{k+1}}}{\text{ext}} h \sum_{j=1}^{m} b_j L\left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)\right).$$

Requiring the curve $\tilde{q}_n(t)$ to be a stationary point of the discretized action provides $n - 2$ internal stage conditions:

$$(3) \quad h \sum_{j=1}^{m} b_j \left(\frac{\partial L}{\partial q}\left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)\right) \phi_p(c_j h) + \frac{\partial L}{\partial \dot{q}}\left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)\right) \dot{\phi}_p(c_j h)\right) = 0, \quad p = 2, ..., n-1.$$

Combining these internal stage conditions with the discrete Euler-Lagrange equations,

$$D_1 L_d(q_{k-1}, q_k) + D_2 L_d(q_k, q_{k+1}) = 0,$$

and the continuity condition $q_k^1 = q_k$ yields the following set of n nonlinear equations,

$$q_k^1 = q_k,$$

$$h \sum_{j=1}^{m} b_j \left(\frac{\partial L}{\partial q}\left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)\right) \phi_p(c_j h) + \frac{\partial L}{\partial \dot{q}}\left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)\right) \dot{\phi}_p(c_j h)\right) = 0, \quad p = 2, ..., n-1,$$

$$(4) \quad h \sum_{j=1}^{m} b_j \left(\frac{\partial L}{\partial q}\left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)\right) \phi_1(c_j h) + \frac{\partial L}{\partial \dot{q}}\left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)\right) \dot{\phi}_1(c_j h)\right) = p_{k-1},$$

which must be solved at each time step k, and the momentum condition:

$$h \sum_{j=1}^{m} b_j \left(\frac{\partial L}{\partial q}\left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)\right) \phi_n(c_j h) + \frac{\partial L}{\partial \dot{q}}\left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)\right) \dot{\phi}_n(c_j h)\right) = p_k,$$

which defines (4) for the next time step. Evaluating $q_{k+1} = \tilde{q}_n(h)$ defines the next step for the discrete Lagrangian flow map:

$$F_{L_d}(q_{k-1}, q_k) = (q_k, q_{k+1}),$$

and because of the choice of basis functions, this is simply $q_{k+1} = q_k^n$.

2.2.2. *n-Refinement.* As is typical for spectral numerical methods (see, for example, Boyd [4]; Trefethen [25]), convergence for spectral variational integrators is achieved by increasing the dimension of the function space, $\mathbb{M}^n([0,h], Q)$. Furthermore, because the order of the discrete Lagrangian also depends on the order of the quadrature rule $\mathcal{G}$, we must also refine the quadrature rule as we refine n. Hence, for examining convergence, we must also consider the quadrature rule as a function of n, $\mathcal{G}_n$. Because of the dependence on n instead of h, we will often examine the discrete Lagrangian L_d as a function of $Q \times Q \times \mathbb{N}$,

$$L_d(q_k, q_{k+1}, n) = \underset{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_k^1 = q_k, q_k^n = q_{k+1}}}{\text{ext}} \mathcal{G}_n\left(L\left(\tilde{q}_n(t), \dot{\tilde{q}}_n(t)\right)\right) = \underset{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_k^1 = q_k, q_k^n = q_{k+1}}}{\text{ext}} h \sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n(c_{n_j} h), \dot{\tilde{q}}_n(c_{n_j} h)\right),$$

as opposed to the more conventional

$$L_d(q_k, q_{k+1}, h) = \underset{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_k^1 = q_k, q_k^n = q_{k+1}}}{\text{ext}} \mathcal{G}\left(L\left(\tilde{q}_n(t), \dot{\tilde{q}}_n(t)\right)\right) = \underset{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_k^1 = q_k, q_k^n = q_{k+1}}}{\text{ext}} h \sum_{j=1}^{m} b_j L\left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)\right).$$

This type of refinement is the foundation for the exceptional convergence properties of spectral variational integrators.

3. Existence, Uniqueness and Convergence

In this section, we will discuss the major important properties of Galerkin variational integrators and spectral variational integrators. The first will be the existence of unique solutions to the internal stage equations (3) for certain types of Lagrangians. The second is the convergence of the one-step map that results from the Galerkin and spectral variational constructions, which will be shown to be optimal in a certain sense. The third and final is the convergence of continuous approximations to the Euler-Lagrange flow which can easily be constructed from Galerkin and spectral variational integrators, and the behavior of geometric invariants associated with the approximate continuous flow. We will show a number of different convergence results associated with these quantities, which demonstrate that Galerkin and spectral variational integrators can be used to compute continuous approximations to the exact solutions of the Euler-Lagrange equations which have excellent convergence and geometric behavior.

3.1. **Existence and Uniqueness.** In general, demonstrating that there exists a unique solution to the internal stage equations for a spectral variational integrator is difficult, and depends on the properties of the Lagrangian. However, assuming a Lagrangian of the form

$$L\left(q,\dot{q}\right) = \frac{1}{2}\dot{q}^T M\dot{q} - V\left(q\right),$$

it is possible to show the existence and uniqueness of the solutions to the implicit equations for the one-step method under appropriate assumptions.

Theorem 3.1. (Existence and Uniqueness of Solutions to the Internal Stage Equations) *Given a Lagrangian* $L: TQ \to \mathbb{R}$ *of the form*

$$L\left(q,\dot{q}\right) = \frac{1}{2}\dot{q}^T M\dot{q} - V\left(q\right),$$

if ∇V *is Lipschitz continuous,* $b_j > 0$ *for every* j *and* $\sum_{i=1}^m b_j = 1$, *and* M *is symmetric positive-definite, then there exists an interval* $[0,h]$ *where there exists a unique solution to the internal stage equations for a spectral variational integrator.*

Proof. We will consider only the case where $q\left(t\right) \in \mathbb{R}$, but the argument generalizes easily to higher dimensions. To begin, we note that for a Lagrangian of the form,

$$L\left(q,\dot{q}\right) = \frac{1}{2}\dot{q}^T M\dot{q} - V\left(q\right)$$

the equations

$$q_k^1 = q_k,$$

$$h\sum_{j=1}^m b_j \left(\frac{\partial L}{\partial q}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\phi_p\left(c_jh\right) + \frac{\partial L}{\partial\dot{q}}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\dot{\phi}_p\left(c_jh\right)\right) = 0, \qquad p = 2,...,n-1,$$

$$h\sum_{j=1}^m b_j \left(\frac{\partial L}{\partial q}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\phi_1\left(c_jh\right) + \frac{\partial L}{\partial\dot{q}}\left(\tilde{q}_n\left(c_jh\right),\dot{\tilde{q}}_n\left(c_jh\right)\right)\dot{\phi}_1\left(c_jh\right)\right) = p_{k-1},$$

take the form

$$(5) \qquad\qquad Aq^i - f\left(q^i\right) = 0,$$

where q^i is the vector of internal weights, $q^i = \left(q_k^1, q_k^2, ..., q_k^n\right)^T$, A is a matrix with entries defined by

$$(6) \qquad A_{1,1} = 1,$$

$$(7) \qquad A_{1,i} = 0, \qquad\qquad\qquad\qquad\qquad\qquad i = 2,...,n,$$

$$(8) \qquad A_{p,i} = h\sum_{j=1}^m b_j M\dot{\phi}_i\left(c_jh\right)\dot{\phi}_p\left(c_jh\right), \qquad\qquad p = 2,...,n; i = 1,...,n,$$

and f is a vector valued function defined by

$$f\left(q^i\right) = \begin{pmatrix} q_k \\ h\sum_{j=1}^m b_j \nabla V\left(\sum_{i=1}^n q_k^i \phi_i\left(c_j h\right)\right)\phi_2 \\ \vdots \\ h\sum_{j=1}^m b_j \nabla V\left(\sum_{i=1}^n q_k^i \phi_i\left(c_j h\right)\right)\phi_{n-1} \\ p_{k-1} \end{pmatrix}.$$

It is important to note that the entries of A depend on h. For now we will assume A is invertible, and that $\left\|A^{-1}\right\| < \left\|A_1^{-1}\right\|$, for where A_1 is the matrix A generated on the interval $[0,1]$. Of course, the properties of A depend on the choice of basis functions $\{\phi_i\}_{i=1}^n$, but we will establish these properties for the polynomial basis later. Defining the map:

$$\Phi\left(q^i\right) = A^{-1}f\left(q^i\right),$$

it is easily seen that (5) is satisfied if and only if $q^i = \Phi\left(q^i\right)$, that is, q^i is a fixed point of $\Phi\left(\cdot\right)$. If we establish that $\Phi\left(\cdot\right)$ is a contraction mapping,

$$\left\|\Phi\left(w^i\right) - \Phi\left(v^i\right)\right\|_\infty \le k\left\|w^i - v^i\right\|_\infty,$$

for some $k < 1$, we can establish the existence of a unique fixed point, and thus show that the steps of the one step method are well-defined. Here, and throughout this section, we use $\left\|\cdot\right\|_p$ to denote the vector or matrix p-norm, as appropriate.

To show that $\Phi\left(\cdot\right)$ is a contraction mapping, we consider arbitrary w^i and v^i:

$$\begin{aligned}
\left\|\Phi\left(w^i\right) - \Phi\left(v^i\right)\right\|_\infty &= \left\|A^{-1}f\left(w^i\right) - A^{-1}f\left(v^i\right)\right\|_\infty \\
&= \left\|A^{-1}\left(f\left(w^i\right) - f\left(v^i\right)\right)\right\|_\infty \\
&\le \left\|A^{-1}\right\|_\infty \left\|f\left(w^i\right) - f\left(v^i\right)\right\|_\infty.
\end{aligned}$$

Considering $\left\|f\left(w^i\right) - f\left(v^i\right)\right\|_\infty$, we see that

$$(9) \qquad \left\|f\left(w^i\right) - f\left(v^i\right)\right\|_\infty = \left|\sum_{j=1}^m b_j\left[\nabla V\left(\sum_{i=1}^n w_k^i\phi_i\left(c_j h\right)\right) - \nabla V\left(\sum_{i=1}^n v_k^i\phi_i\left(c_j h\right)\right)\right]\phi_{p^*}\left(c_j h\right)\right|,$$

for some appropriate index p^*. Note that the first and last terms of $\left\|f\left(w^i\right) - f\left(v^i\right)\right\|_\infty$ will vanish, so the maximum element must take the form of (9). Let $\phi^i\left(t\right) = \left(\phi_1\left(t\right), \phi_2\left(t\right), ..., \phi_n\left(t\right)\right)$. Let C_L be the Lipschitz constant for $\nabla V\left(q\right)$. Now

$$\begin{aligned}
\left\|f\left(w^i\right) - f\left(v^i\right)\right\|_\infty &= \left|h\sum_{j=1}^m b_j\left[\nabla V\left(\sum_{i=1}^n w_k^i\phi_i\left(c_j h\right)\right) - \nabla V\left(\sum_{i=1}^n v_k^i\phi_i\left(c_j h\right)\right)\right]\phi_{p^*}\left(c_j h\right)\right| \\
&\le h\sum_{j=1}^m |b_j|\left|\left[\nabla V\left(\sum_{i=1}^n w_k^i\phi_i\left(c_j h\right)\right) - \nabla V\left(\sum_{i=1}^n v_k^i\phi_i\left(c_j h\right)\right)\right]\right| \left|\phi_{p^*}\left(c_j h\right)\right| \\
&\le h\sum_{j=1}^m b_j C_L\left|\sum_{i=1}^n w_k^i\phi_i\left(c_j h\right) - \sum_{i=1}^n v_k^i\phi_i\left(c_j h\right)\right| \left|\phi_{p^*}\left(c_j h\right)\right| \\
&= h\sum_{j=1}^m b_j C_L\left|\sum_{i=1}^n \left(w_k^i - v_k^i\right)\phi_i\left(c_j h\right)\right| \left|\phi_{p^*}\left(c_j h\right)\right| \\
&\le h\sum_{j=1}^m b_j C_L\left\|w^i - v^i\right\|_\infty \left\|\phi^i\left(c_j h\right)\right\|_1 \left|\phi_{p^*}\left(c_j h\right)\right| \\
&\le h\sum_{j=1}^m b_j C_L \max_j\left(\left\|\phi^i\left(c_j h\right)\right\|_1 \left|\phi_{p^*}\left(c_j h\right)\right|\right)\left\|w^i - v^i\right\|_\infty
\end{aligned}$$

$$= hC_L \max_j \left(\left\| \phi^i \left(c_j h \right) \right\|_1 \left| \phi_{p^*} \left(c_j h \right) \right| \right) \left\| w^i - v^i \right\|_\infty .$$

Hence, we derive the inequality

$$\left\| \Phi \left(w^i \right) - \Phi \left(v^i \right) \right\|_\infty \leq h \left\| A^{-1} \right\|_\infty C_L \max_j \left(\left\| \phi^i \left(c_j h \right) \right\|_1 \left| \phi_{p^*} \left(c_j h \right) \right| \right) \left\| w^i - v^i \right\|_\infty ,$$

and since by assumption $\left\| A^{-1} \right\|_\infty \leq \left\| A_1^{-1} \right\|_\infty$,

$$\left\| \Phi \left(w^i \right) - \Phi \left(v^i \right) \right\|_\infty \leq h \left\| A_1^{-1} \right\|_\infty C_L \max_j \left(\left\| \phi^i \left(c_j h \right) \right\|_1 \left| \phi_{p^*} \left(c_j h \right) \right| \right) \left\| w^i - v^i \right\|_\infty .$$

Thus if:

$$h < \left(\left\| A_1^{-1} \right\|_\infty C_L \max_j \left(\left\| \phi^i \left(c_j h \right) \right\|_1 \left| \phi_{p^*} \left(c_j h \right) \right| \right) \right)^{-1} ,$$

then

$$\left\| \Phi \left(w^i \right) - \Phi \left(v^i \right) \right\|_\infty \leq k \left\| w^i - v^i \right\|_\infty ,$$

where $k < 1$, which establishes that $\Phi \left(\cdot \right)$ is a contraction mapping, and establishes the existence of a unique fixed point, and thus the existence of unique steps of the one step method. $\qquad \square$

A critical assumption made during the proof of existence and uniqueness is that the matrix A is nonsingular. This property depends on the choice of basis functions ϕ_i. However, using a polynomial basis, like Lagrange interpolation polynomials, it can be shown that A is invertible.

Lemma 3.1. (A is invertible) *If $\{\phi_i\}_{i=1}^n$ is a polynomial basis of P_n, the space of polynomials of degree at most n, M is symmetric positive-definite, and the quadrature rule is order at least $2n + 1$, then A defined by (6) – (8) is invertible.*

Proof. We begin by considering the equation:

$$Aq^i = 0.$$

Let $\tilde{q}_n \left(t \right) = \sum_{i=1}^n q_k^i \phi_i \left(t \right)$. Considering the definition of A, $Aq^i = 0$ holds if and only if the following equations hold:

$$\tilde{q}_n \left(0 \right) = 0,$$

(10)
$$h \sum_{j=1}^m b_j M \dot{\tilde{q}}_n \left(c_j h \right) \dot{\phi}_p \left(c_j h \right) = 0, \qquad\qquad p = 1, ..., (n - 1).$$

It can easily be seen that $\left\{ \dot{\phi}_i \right\}_{i=1}^{n-1}$ is a basis of P_{n-1}. Using the assumption that the quadrature rule is of order at least $2n - 1$ and that M is symmetric positive-definite, we can see that (10) implies:

$$\int_0^h M \dot{\tilde{q}}_n \left(t \right) \dot{\phi}_p \left(t \right) \mathrm{d}t = 0, \qquad\qquad p = 1, ..., (n - 1),$$

but,

$$\int_0^h M \dot{\tilde{q}}_n \left(t \right) \dot{\phi}_i \left(t \right) \mathrm{d}t = 0$$

implies

$$\left\langle \dot{\tilde{q}}_n, \dot{\phi}_p \right\rangle = 0,$$

where $\langle \cdot, \cdot \rangle$ is the standard L^2 inner product on $[0, h]$. Since $\left\{ \dot{\phi}_i \right\}_{i=1}^{n-1}$ forms a basis for P_{n-1}, $\dot{\tilde{q}}_n \in P_{n-1}$, and $\langle \cdot, \cdot \rangle$ is non-degenerate, this implies that $\dot{\tilde{q}}_n \left(t \right) = 0$. Thus,

$$\tilde{q}_n \left(0 \right) = 0$$
$$\dot{\tilde{q}}_n \left(t \right) = 0$$

which implies that $\tilde{q}_n(t) = 0$ and hence $q^i = 0$. Thus, $Aq^i = 0$ then $q^i = 0$, from which it follows that A is non-singular. $\qquad\square$

Another subtle difficulty is that the matrix A is a function of h. Since we assumed that $\left\|A^{-1}\right\|_\infty$ is bounded to prove Theorem 3.1, we must show that for any choice of h, the quantity $\left\|A^{-1}\right\|_\infty$ is bounded. We will do this by establishing $\left\|A^{-1}\right\|_\infty \le \left\|A_1^{-1}\right\|_\infty$, where A_1 is A defined with $h = 1$. By Lemma 3.1, we know that $\left\|A_1^{-1}\right\|_\infty < \infty$, which establishes the upper bound for $\left\|A^{-1}\right\|_\infty$. This argument is easily generalized for a higher upper bound on h.

Lemma 3.2. $(\left\|A^{-1}\right\|_\infty \le \left\|A_1^{-1}\right\|_\infty)$ *For the matrix A defined by (6) – (8), if $h < 1$, $\left\|A^{-1}\right\|_\infty < \left\|A_1^{-1}\right\|_\infty$ where A_1 is A defined on the interval $[0, 1]$.*

Proof. We begin the proof by examining how A changes as a function of h. First, let $\{\phi_i\}_{i=1}^n$ be the basis for the interval $[0, 1]$. Then for the interval $[0, h]$, the basis functions are

$$\phi_i^h(t) = \phi_i\left(\frac{t}{h}\right)$$

and hence the derivatives are:

$$\dot{\phi}_i^h(t) = \frac{1}{h}\dot{\phi}_i\left(\frac{t}{h}\right).$$

Thus, if A_1 is the matrix defined by (6) – (8) on the interval $[0, 1]$, then for the interval $[0, h]$,

$$A = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{h}I_{(n-1)\times(n-1)} \end{pmatrix} A_1,$$

where $I_{n\times n}$ is the $n \times n$ identity matrix. This gives

$$A^{-1} = A_1^{-1}\begin{pmatrix} 1 & 0 \\ 0 & hI_{(n-1)\times(n-1)} \end{pmatrix}$$

which gives

$$\left\|A^{-1}\right\|_\infty = \left\|A_1^{-1}\begin{pmatrix} 1 & 0 \\ 0 & hI_{(n-1)\times(n-1)} \end{pmatrix}\right\|_\infty \le \left\|A_1^{-1}\right\|_\infty \left\|\begin{pmatrix} 1 & 0 \\ 0 & hI_{(n-1)\times(n-1)} \end{pmatrix}\right\|_\infty = \left\|A_1^{-1}\right\|_\infty,$$

which proves the statement. $\qquad\square$

3.2. **Order Optimal and Geometric Convergence.** To determine the rate of convergence for spectral variational integrators, we will utilize Theorem 1.1 and a simple extension of Theorem 1.1:

Theorem 3.2. (Extension of Theorem 1.1 to Geometric Convergence) *Given a regular Lagrangian L and corresponding Hamiltonian H, the following are equivalent for a discrete Lagrangian $L_d(q_0, q_1, n)$:*

(1) *there exists a positive constant K, where $K < 1$, such that the discrete Hamiltonian map for L_d has error $\mathcal{O}(K^n)$,*

(2) *there exists a positive constant K, where $K < 1$, such that the discrete Legendre transforms of L_d have error $\mathcal{O}(K^n)$,*

(3) *there exists a positive constant K, where $K < 1$, such that L_d is equivalent to a discrete Lagrangian with error $\mathcal{O}(K^n)$.*

This theorem provides a fundamental tool for the analysis of Galerkin variational methods. Its proof is almost identical to that of Theorem 1.1, and can be found in the appendix. The critical result is that the order of the error of the discrete Hamiltonian flow map, from which we construct the discrete flow, has the same order as the discrete Lagrangian from which it is constructed. Thus, in order to determine the order of the error of the flow generated by spectral variational integrators, we need only determine how well the discrete Lagrangian approximates the exact discrete Lagrangian. This is a key result which greatly reduces the difficulty of the error analysis of Galerkin variational integrators.

Naturally, the goal of constructing spectral variational integrators is constructing a variational method that has geometric convergence. To this end, it is essential to establish that Galerkin type integrators inherit the convergence properties of the spaces which are used to construct them. The order optimality

result is related to the problem of Γ-convergence (see, for example, Dal Maso [6]), as the Galerkin discrete Lagrangians are given by extremizers of an approximating sequence of variational problems, and the exact discrete Lagrangian is the extremizer of the limiting variational problem. The Γ-convergence of variational integrators was studied in Müller and Ortiz [22], and our approach involves a refinement of their analysis. We now state our results, which establish not only the geometric convergence of spectral variational integrators, but also order optimality of all Galerkin variational integrators under appropriate smoothness assumptions.

Theorem 3.3. (Order Optimality of Galerkin Variational Integrators) *Given an interval $[0, h]$ and a Lagrangian $L : TQ \to \mathbb{R}$, let $\bar{q}$ be the exact solution to the Euler-Lagrange equations subject to the conditions $\bar{q}(0) = q_0$ and $\bar{q}(h) = q_h$, and let $\tilde{q}_n$ be the stationary point of a Galerkin variational discrete action, i.e. if $L_d^G : Q \times Q \times \mathbb{R} \to \mathbb{R}$,*

$$L_d^G(q_0, q_h, h) = \operatorname*{ext}_{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_h}} \mathbb{S}_d(\{q_i\}_{i=1}^n) = \operatorname*{ext}_{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_h}} h \sum_{j=1}^m b_j L(q_n(c_j h), \dot{q}_n(c_j h)),$$

then

$$\tilde{q}_n = \operatorname*{argmin}_{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_h}} h \sum_{j=1}^m b_j L(q_n(c_j h), \dot{q}_n(c_j h)).$$

If:

(1) *there exists a constant C_A independent of h, such that, for each h, there exists a curve $\hat{q}_n \in \mathbb{M}^n([0, h], Q)$, such that,*

$$\left| \left(\hat{q}_n(t), \dot{\hat{q}}_n(t) \right) - (\bar{q}(t), \dot{\bar{q}}(t)) \right| \leq C_A h^n,$$

(2) *there exists a closed and bounded neighborhood $U \subset TQ$, such that $(\bar{q}(t), \dot{\bar{q}}(t)) \in U$, $\left(\hat{q}_n(t), \dot{\hat{q}}_n(t) \right) \in U$ for all t, and all partial derivatives of L are continuous on U,*

(3) *for the quadrature rule $\mathcal{G}(f) = h \sum_{j=1}^m b_j f(c_j h) \approx \int_0^h f(t) \, dt$, there exists a constant C_g, such that,*

$$\left| \int_0^h L(q_n(t), \dot{q}_n(t)) \, dt - h \sum_{j=1}^m b_j L(q_n(c_j h), \dot{q}_n(c_j h)) \right| \leq C_g h^{n+1},$$

for any $q_n \in \mathbb{M}^n([0, h], Q)$,

(4) *and the stationary points $\bar{q}$, $\tilde{q}_n$ minimize their respective actions,*

then

$$\left| L_d^E(q_0, q_h, h) - L_d^G(q_0, q_h, h) \right| \leq C_{op} h^{n+1},$$

for some constant C_{op} independent of h, i.e. discrete Lagrangian L_d has error $\mathcal{O}(h^{n+1})$, and hence the discrete Hamiltonian flow map has error $\mathcal{O}(h^{n+1})$.

Proof. First, we rewrite both the exact discrete Lagrangian and the Galerkin discrete Lagrangian:

$$\left| L_d^E(q_0, q_h, h) - L_d^G(q_0, q_h, h) \right| = \left| \int_0^h L(\bar{q}(t), \dot{\bar{q}}(t)) \, dt - \mathcal{G}\left(L(\tilde{q}_n(t), \dot{\tilde{q}}_n(t)) \right) \right|$$

$$= \left| \int_0^h L(\bar{q}(t), \dot{\bar{q}}(t)) \, dt - h \sum_{j=1}^m b_j L(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)) \right|$$

$$= \left| \int_0^h L(\bar{q}, \dot{\bar{q}}) \, dt - h \sum_{j=1}^m b_j L(\tilde{q}_n, \dot{\tilde{q}}_n) \right|,$$

10

where in the last line, we have suppressed the t argument, a convention we will continue throughout the proof. Now we introduce the action evaluated on the $\hat{q}_n$ curve, which is an approximation with error $\mathcal{O}\left(h^n\right)$ to the exact solution $\bar{q}$:

$$\left| \int_0^h L\left(\bar{q}, \dot{\bar{q}}\right) \mathrm{d}t - h \sum_{j=1}^m b_j L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \right| = \left| \int_0^h L\left(\bar{q}, \dot{\bar{q}}\right) \mathrm{d}t - \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t \right.$$

$$\left. + \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t - h \sum_{j=1}^m b_j L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \right|$$

(11a)
$$\leq \left| \int_0^h L\left(\bar{q}, \dot{\bar{q}}\right) \mathrm{d}t - \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t \right|$$

(11b)
$$+ \left| \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t - h \sum_{j=1}^m b_j L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \right|.$$

Considering the first term (11a):

$$\left| \int_0^h L\left(\bar{q}, \dot{\bar{q}}\right) \mathrm{d}t - \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t \right| = \left| \int_0^h L\left(\bar{q}, \dot{\bar{q}}\right) - L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t \right|$$

$$\leq \int_0^h \left| L\left(\bar{q}, \dot{\bar{q}}\right) - L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \right| \mathrm{d}t.$$

By assumption, all partials of L are continuous on U, and since U is closed and bounded, this implies L is Lipschitz on U. Let L_α denote that Lipschitz constant. Since, again by assumption, $\left(\bar{q}, \dot{\bar{q}}\right) \in U$ and $\left(\hat{q}_n, \dot{\hat{q}}_n\right) \in U$, we can rewrite:

$$\int_0^h \left| L\left(\bar{q}, \dot{\bar{q}}\right) - L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \right| \mathrm{d}t \leq \int_0^h L_\alpha \left| \left(\bar{q}, \dot{\bar{q}}\right) - \left(\hat{q}_n, \dot{\hat{q}}_n\right) \right| \mathrm{d}t$$

$$\leq \int_0^h L_\alpha C_A h^n \mathrm{d}t$$

$$= L_\alpha C_A h^{n+1},$$

where we have made use of the best approximation estimate. Hence,

(12)
$$\left| \int_0^h L\left(\bar{q}, \dot{\bar{q}}\right) \mathrm{d}t - \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t \right| \leq L_\alpha C_1 h^{n+1}.$$

Next, considering the second term (11b),

$$\left| \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t - h \sum_{j=1}^m b_j L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \right|,$$

since $\tilde{q}_n$, the stationary point of the discrete action, minimizes its action and $\hat{q}_n \in \mathbb{M}^n\left([0, h], Q\right)$,

(13)
$$h \sum_{j=1}^m b_j L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \leq h \sum_{j=1}^m b_j L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \leq \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t + C_g h^{n+1}$$

where the inequalities follow from the assumptions on the order of the quadrature rule. Furthermore,

$$h \sum_{j=1}^m b_j L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \geq \int_0^h L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \mathrm{d}t - C_g h^{n+1}$$

$$\geq \int_0^h L\left(\bar{q}, \dot{\bar{q}}\right) \mathrm{d}t - C_g h^{n+1}$$

11

(14)
$$\geq \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t - L_\alpha C_A h^{n+1} - C_g h^{n+1},$$

where the inequalities follow from (12), the order of the quadrature rule, and the assumption that $\bar{q}$ minimizes its action. Putting (13) and (14) together, we can conclude:

(15)
$$\left| \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t - h \sum_{j=1}^m b_j L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \right| \leq (L_\alpha C_A + C_g) \, h^{n+1}.$$

Now, combining the bounds (12) and (15) in (11a) and (11b), we can conclude

$$\left| L_d^E\left(q_0, q_h, h\right) - L_d^G\left(q_0, q_h, h\right) \right| \leq (2L_\alpha C_A + C_g) \, h^{n+1}$$

which, combined with Theorem 1.1, establishes the order of the error of the integrator. $\qquad\square$

The above proof establishes a significant convergence result for Galerkin variational integrators, namely that for sufficiently well behaved Lagrangians, Galerkin variational integrators will produce discrete approximate flows that converge to the exact flow as $h \to 0$ with the highest possible order allowed by the approximation space, provided the quadrature rule is of sufficiently high order.

We will discuss assumption 4 in §3.3. While in general we cannot assume that stationary points of the action are minimizers, it can be shown that for Lagrangians of the canonical form

$$L\left(q, \dot{q}\right) = \dot{q}^T M \dot{q} - V\left(q\right),$$

under some mild assumptions on the derivatives of V and the accuracy of the quadrature rule, there always exists an interval $[0, h]$ over which stationary points are minimizers. In §3.3 we will show the result extends to the discretized action of Galerkin variational integrators. A similar result was established in Müller and Ortiz [22].

Geometric convergence of spectral variational integrators is not strictly covered under the proof of order optimality. While the above theorem establishes convergence of Galerkin variational integrators by shrinking h, the interval length of each discrete Lagrangian, spectral variational integrators achieve convergence by holding the interval length of each discrete Lagrangian constant and increasing the dimension of the approximation space $\mathbb{M}^n\left([0, h], Q\right)$. Thus, for spectral variational integrators, we have the following analogous convergence theorem:

Theorem 3.4. (Geometric Convergence of Spectral Variational Integrators) *Given an interval $[0, h]$ and a Lagrangian $L : TQ \to \mathbb{R}$, let $\bar{q}$ be the exact solution to the Euler-Lagrange equations, and $\tilde{q}_n$ be the stationary point of the spectral variational discrete action:*

$$L_d^S\left(q_0, q_h, n\right) = \operatorname*{ext}_{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_h}} \mathbb{S}_d\left(\{q_i\}_{i=1}^n\right) = \operatorname*{ext}_{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_h}} h \sum_{j=0}^{m_n} b_{n_j} L\left(q_n\left(c_{n_j} h\right), \dot{q}_n\left(c_{n_j} h\right)\right).$$

If:

(1) *there exists constants C_A, K_A, $K_A < 1$, independent of n such that, for each n, there exists a curve $\hat{q}_n \in \mathbb{M}^n\left([0, h], Q\right)$, such that,*

$$\left| \left(\bar{q}, \dot{\bar{q}}\right) - \left(\hat{q}_n, \dot{\hat{q}}_n\right) \right| \leq C_A K_A^n,$$

(2) *there exists a closed and bounded neighborhood $U \subset TQ$, such that, $\left(\bar{q}\left(t\right), \dot{\bar{q}}\left(t\right)\right) \in U$ and $\left(\hat{q}_n\left(t\right), \dot{\hat{q}}_n\left(t\right)\right) \in U$ for all t and n, and all partial derivatives of L are continuous on U,*

(3) *for the sequence of quadrature rules $\mathcal{G}_n\left(f\right) = \sum_{j=1}^{m_n} b_{n_j} f\left(c_{n_j} h\right) \approx \int_0^h f\left(t\right) dt$, there exists constants C_g, K_g, $K_g < 1$, independent of n such that*

$$\left| \int_0^h L\left(q_n\left(t\right), \dot{q}_n\left(t\right)\right) dt - h \sum_{j=1}^{m_n} b_{n_j} L\left(q_n\left(c_{n_j} h\right), \dot{q}_n\left(c_{n_j} h\right)\right) \right| \leq C_g K_g^n,$$

for any $q_n \in \mathbb{M}^n\left([0, h], Q\right)$,

(4) *and the stationary points $\bar{q}, \tilde{q}_n$ minimize their respective actions,*

then

(16)
$$\left| L_d^E\left(q_0, q_1\right) - L_d^S\left(q_0, q_1, n\right) \right| \leq C_s K_s^n$$

for some constants C_s, K_s, $K_s < 1$, independent of n, and hence the discrete Hamiltonian flow map has error $\mathcal{O}\left(K_s^n\right)$.

The proof of the above theorem is very similar to that of order optimality, and would be tedious to repeat here. It can be found in the appendix. The main differences between the proofs are the assumption of the sequence of converging functions in the increasingly high-dimensional approximation spaces, and the assumption of a sequence of increasingly high-order quadrature rules. These assumptions are used in the obvious way in the modified proof.

3.3. **Minimization of the Action.** One of the major assumptions made in the convergence theorems (3.3) and (3.4) is that the the stationary points of both the continuous and discrete actions are minimizers over the interval $[0, h]$. This type of minimization requirement is similar to the one made in the paper on Γ-convergence of variational integrators by Müller and Ortiz [22]. In fact, the results in Müller and Ortiz [22] can easily be extended to demonstrate that for sufficiently well-behaved Lagrangians of the form

$$L\left(q, \dot{q}\right) = \frac{1}{2}\dot{q}^T M \dot{q} - V\left(q\right),$$

where $q \in C^2\left([0, h], Q\right)$, there exists an interval $[0, h]$, such that stationary points of the Galerkin action are minimizers.

Theorem 3.5. *Consider a Lagrangian of the form*

$$L\left(q, \dot{q}\right) = \frac{1}{2}\dot{q}^T M \dot{q} - V\left(q\right)$$

where $q \in C^2\left([0, h], Q\right)$ and each component q^d of q, $q^d \in C^2\left([0, h], Q\right)$, is a polynomial of degree at most s. Assume M is symmetric positive-definite and all second-order partial derivatives of V exist, and are continuous and bounded. Then, there exists a time interval $[0, h]$ such that stationary points of the discrete action,

$$\mathbb{S}_d\left(\left\{q_k^i\right\}_{i=1}^n\right) = h\sum_{j=1}^m b_j\left(\frac{1}{2}\dot{\tilde{q}}_n\left(c_j h\right)^T M \dot{\tilde{q}}_n\left(c_j h\right) - V\left(\tilde{q}_n\left(c_j h\right)\right)\right),$$

on this time interval are minimizers if the quadrature rule used to construct the discrete action is of order at least $2s + 1$.

We quickly note that the assumption that each component of q, q^d, is a polynomial of degree *at most s* allows for discretizations where different components of the configuration space are discretized with polynomials of different degrees. This allows for more efficient discretizations where slower evolving components are discretized with lower-degree polynomials than faster evolving ones.

Proof. Let $\tilde{q}_n$ be a stationary point of the discrete action $\mathbb{S}_d\left(\cdot\right)$, and let δq be an arbitrary perturbation of the stationary point $\tilde{q}_n$, under the conditions $\delta q^d \in P_{S_d}$, $\delta q\left(0\right) = \delta q\left(h\right) = 0$, which is uniquely defined by $\left\{\delta q_k^i\right\}_{i=1}^n \subset Q$. Then,

$$\mathbb{S}_d\left(\left\{q_k^i + \delta q_k^i\right\}_{i=1}^n\right) - \mathbb{S}_d\left(\left\{q_k^i\right\}_{i=1}^n\right)$$

$$= h\sum_j^m b_j\left(\frac{1}{2}\left(\dot{\tilde{q}}_n + \delta\dot{q}\right)^T M\left(\dot{\tilde{q}}_n + \delta\dot{q}\right) - V\left(\tilde{q}_n + \delta q\right)\right) - h\sum_j^m b_j\left(\frac{1}{2}\dot{\tilde{q}}_n^T M \dot{\tilde{q}}_n - V\left(\tilde{q}_n\right)\right)$$

$$= h\sum_j^m b_j\left(\frac{1}{2}\left(\dot{\tilde{q}}_n + \delta\dot{q}\right)^T M\left(\dot{\tilde{q}}_n + \delta\dot{q}\right) - V\left(\tilde{q}_n + \delta q\right) - \frac{1}{2}\dot{\tilde{q}}_n^T M \dot{\tilde{q}}_n + V\left(\tilde{q}_n\right)\right).$$

Making use of Taylor's remainder theorem, we expand:

$$V\left(\tilde{q}_n + \delta q\right) = V\left(\tilde{q}_n\right) + \nabla V\left(\tilde{q}_n\right) \cdot \delta q + \frac{1}{2}\delta\tilde{q}_n^T R \delta\tilde{q}_n,$$

13

where $|R_{lm}| \leq \sup_{l,m} \left| \frac{\partial^2 V}{\partial q_l \partial q_m} \right|$. Using this expansion, we rewrite

$$\mathbb{S}_d \left(\{q_k^i + \delta q_k^i\}_{i=1}^n \right) - \mathbb{S}_d \left(\{q_k^i\}_{i=1}^n \right) = h \sum_j^m b_j \left(\frac{1}{2} \left(\dot{\tilde{q}}_n + \delta \dot{q} \right)^T M \left(\dot{\tilde{q}}_n + \delta \dot{q} \right) - V \left(\tilde{q}_n \right) - \nabla V \left(\tilde{q}_n \right) \cdot \delta q \right.$$

$$\left. - \frac{1}{2} \delta q^T R \delta q - \left(\frac{1}{2} \dot{\tilde{q}}_n^T M \dot{\tilde{q}}_n + V \left(\tilde{q}_n \right) \right) \right)$$

which, given the symmetry in M, rearranges to:

$$\mathbb{S}_d \left(\{q_k^i + \delta q_k^i\}_{i=1}^n \right) - \mathbb{S}_d \left(\{q_k^i\}_{i=1}^n \right) = h \sum_j^m b_j \left(\dot{\tilde{q}}_n^T M \delta \dot{q} - \nabla V \left(\tilde{q}_n \right) \cdot \delta q + \frac{1}{2} \delta \dot{q}^T M \delta \dot{q} - \frac{1}{2} \delta q^T R \delta q \right).$$

Now, it should be noted that the stationarity condition for the discrete Euler-Lagrange equations is

$$h \sum_{j=1}^m b_j \left(\dot{\tilde{q}}_n^T M \delta \dot{q} - \nabla V \left(\tilde{q}_n \right) \cdot \delta q \right) = 0$$

for arbitrary δq, which allows us to simplify the expression to

$$\mathbb{S}_d \left(\{q_k^i + \delta q_k^i\}_{i=1}^n \right) - \mathbb{S}_d \left(\{q_k^i\}_{i=1}^n \right) = h \sum_j^m b_j \left(\frac{1}{2} \delta \dot{q}^T M \delta \dot{q} - \frac{1}{2} \delta q^T R \delta q \right).$$

Now, using the assumption that the partial derivatives of V are bounded, $|R_{lm}| \leq \left| \frac{\partial^2 V}{\partial q_l \partial q_m} \right| < C_R$, and standard matrix inequalities, we get the inequality:

(17) $$\delta q^T R \delta q \leq \|R \delta q\|_2 \|\delta q\|_2 \leq \|R\|_2 \|\delta q\|_2^2 \leq \|R\|_F \|\delta q\|_2^2 \leq D C_R \|\delta q\|_2^2 = D C_R \delta q^T \delta q,$$

where D is the number of spatial dimensions of Q. Thus

$$h \sum_j^m b_j \left(\frac{1}{2} \delta \dot{q}^T M \delta \dot{q} - \frac{1}{2} \delta q^T R \delta q \right) \geq h \sum_j^m b_j \left(\frac{1}{2} \delta \dot{q}^T M \delta \dot{q} - \frac{1}{2} D C_R \delta q^T \delta q \right).$$

Because M is symmetric positive-definite, there exists $m > 0$ such that $x^T M x \geq m x^T x$ for any x. Hence,

$$h \sum_j^m b_j \left(\frac{1}{2} \delta \dot{q}^T M \delta \dot{q} - \frac{1}{2} D C_R \delta q^T \delta q \right) \geq h \sum_j^m b_j \left(\frac{1}{2} m \delta \dot{q}^T \delta \dot{q} - \frac{1}{2} D C_R \delta q^T \delta q \right).$$

Now, we note that since each component of δq is a polynomial of degree at most s, $\delta q^T \delta q$ and $\delta \dot{q}^T \delta \dot{q}$ are both polynomials of degree less than or equal to $2s$. Since our quadrature rule is of order $2s + 1$, the quadrature rule is exact, and we can rewrite

$$h \sum_j^m b_j \left(\frac{1}{2} m \delta \dot{q}^T \delta \dot{q} - \frac{1}{2} D C_R \delta q^T \delta q \right) = \frac{1}{2} \int_0^h m \delta \dot{q}^T \delta \dot{q} - D C_R \delta q^T \delta q \mathrm{dt}$$

$$= \frac{1}{2} \left(\int_0^h m \delta \dot{q}^T \delta \dot{q} \mathrm{dt} - \int_0^h D C_R \delta q^T \delta q \mathrm{dt} \right).$$

From here, we note that $\delta q \in H_0^1 \left([0, h], Q \right)$, and make use of the Poincaré inequality to conclude

$$\frac{1}{2} \left(\int_0^h m \delta \dot{q}^T \delta \dot{q} \mathrm{dt} - \int_0^h n C_R \delta q^T \delta q \mathrm{dt} \right) \geq \frac{1}{2} \left(m \frac{\pi^2}{h^2} \int_0^h \delta q^T \delta q \mathrm{dt} - D C_R \int_0^h \delta q^T \delta q \mathrm{dt} \right)$$

$$= \frac{1}{2} \left(\frac{m \pi^2}{h^2} - D C_R \right) \int_0^h \delta q^T \delta q \mathrm{dt}.$$

Since $\int_0^h \delta q^T \delta q \mathrm{dt} > 0$,

$$\mathbb{S}_d \left(\{q_k^i + \delta q_k^i\}_{i=1}^n \right) - \mathbb{S}_d \left(\{q_k^i\}_{i=1}^n \right) \geq \frac{1}{2} \left(\frac{m \pi^2}{h^2} - D C_R \right) \int_0^h \delta q^T \delta q \mathrm{dt} > 0$$

so long as $h < \sqrt{\frac{m\pi^2}{DC_R}}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ □

3.4. Convergence of Galerkin Curves and Noether Quantities.

3.4.1. *Galerkin Curves.* In order to construct the one-step method, spectral variational integrators determine a curve,

$$\tilde{q}_n(t) = \sum_{i=1}^{n} q_k^i \phi_i(t),$$

which satisfies

$$\tilde{q}_n(t) = \underset{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_n(0)=q_k, q_n(h)=q_{k+1}}}{\text{argmin}} \; h \sum_{j=1}^{m} b_j L\left(\tilde{q}_n(c_j h), \dot{\tilde{q}}_n(c_j h)\right).$$

Evaluating this curve at h defines the next step of the one-step method, $q_{k+1} = \tilde{q}_n(h)$, but the curve itself has many desirable properties which makes it a good continuous approximation to the true solution of the Euler Lagrange equations $\bar{q}(t)$. In this section, we will examine some of the favorable properties of $\tilde{q}_n(t)$, hereafter referred to as the *Galerkin curve.*

However, before discussing the properties of the Galerkin curve, it is useful review the different curves with which we are working. We have already defined the Galerkin curve, $\tilde{q}_n(t)$, and we will also be making use of the local solution to the Euler-Lagrange equations $\bar{q}(t)$, where

$$\bar{q}(t) = \underset{\substack{q \in C^2([0,h],Q) \\ q(0)=q_k, q(h)=q_{k+1}}}{\text{argmin}} \int_0^h L\left(q(t), \dot{q}(t)\right) \mathrm{d}t.$$

However, while for each interval $\bar{q}$ satisfies the Euler-Lagrange equations exactly, it is not the exact solution of the Euler-Lagrange equations globally, as $q_k \neq \Phi_{kh}(q_0, \dot{q}_0)$, where $\Phi_t(q_0, \dot{q}_0)$ is the flow of the Euler-Lagrange vector field. This is particularly important when discussing invariants, where the invariants of $\bar{q}$ remain constant within a time-step, but not from time-step to time-step.

The first property of the Galerkin curve that we will examine is its rate of convergence to the true flow of the Euler-Lagrange vector field. There are two general sources of error that affect the convergence of these curves, the first being the accuracy to which the curves approximate the local solution to the Euler-Lagrange equations over the interval $[0, h]$ with the boundary (q_k, q_{k+1}), and the second being the accuracy of the boundary conditions (q_k, q_{k+1}) as approximations to a true sampling of the exact flow. Numerical experiments will show that often the second source of error dominates the first, causing the Galerkin curves to converge at the same rate as the one-step map. However, the accuracy to which the Galerkin curves approximate the true minimizers independent of the error of the boundary can also be established under appropriate assumptions about the action. Two theorems which establish this convergence are presented below.

Before we state the theorems, we quickly recall the definitions of the *Sobolev Norm* $\|\cdot\|_{W^{1,p}([0,h])}$,

$$\|f\|_{W^{1,p}([0,h])} = \left(\|f\|_{L^p([0,h])}^p + \left\|\dot{f}\right\|_{L^p([0,h])}^p \right)^{\frac{1}{p}} = \left(\int_0^h |f|^p \, \mathrm{d}t + \int_0^h \left|\dot{f}\right|^p \, \mathrm{d}t \right)^{\frac{1}{p}}.$$

We will make extensive use of this norm when examining convergence of Galerkin curves.

Theorem 3.6. (Geometric Convergence of Galerkin Curves with n-Refinement) *Under the same assumptions as Theorem 3.4, if at $\bar{q}$, the action is twice Frechet differentiable, and if the second Frechet derivative of the action $D^2\mathfrak{S}(\cdot)[\cdot,\cdot]$ is coercive in a neighborhood U of $\bar{q}$, that is,*

$$D^2\mathfrak{S}(\nu)[\delta q, \delta q] \geq C_f \|\delta q\|_{W^{1,1}([0,h])}^2,$$

for all curves $\delta q \in H_0^1([0,h], Q)$ and all $\nu \in U$, then the curves which minimize the discrete action converge to the true solution geometrically with n-refinement with respect to $\|\cdot\|_{W^{1,1}([0,h])}$. Specifically, if the discrete Hamiltonian flow map has error $\mathcal{O}(K_s^n)$, $K_s < 1$, then the Galerkin curves have error $\mathcal{O}\left(\sqrt{K_s}^n\right)$.

Proof. We start with the bound (16) given at the end of Theorem 3.4,

$$\left| L_d^E \left(q_k, q_{k+1} \right) - L_d^S \left(q_k, q_{k+1}, n \right) \right| \le C_s K_s^n$$

and expand using the definitions of $L_d^E \left(q_k, q_{k+1} \right)$ and $L_d^S \left(q_k, q_{k+1}, n \right)$, as well as the assumed accuracy of the quadrature rule $\mathcal{G}_n$ to derive

$$(18) \qquad C_s K_s^n \ge \left| L_d^E \left(q_k, q_{k+1} \right) - L_d^S \left(q_k, q_{k+1}, n \right) \right|$$

$$= \left| \int_0^h L \left(\bar{q}, \dot{\bar{q}} \right) \mathrm{d}t - h \sum_{j=1}^{m_n} b_{n_j} L \left(\tilde{q}_n \left(c_{n_j} h \right), \dot{\tilde{q}}_n \left(c_{n_j} h \right) \right) \right|$$

$$(19) \qquad \ge \left| \int_0^h L \left(\bar{q}, \dot{\bar{q}} \right) \mathrm{d}t - \int_0^h L \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) \mathrm{d}t \right| - C_g K_g^n$$

$$= \left| \mathfrak{S} \left(\tilde{q}_n \right) - \mathfrak{S} \left(\bar{q} \right) \right| - C_g K_g^n$$

which implies:

$$\left(C_s + C_g \right) K_s^n \ge \left| \mathfrak{S} \left(\tilde{q}_n \right) - \mathfrak{S} \left(\bar{q} \right) \right|$$

because $K_s \ge K_g$, (see the proof of Theorem 3.4 in the appendix). Using this inequality, we make use of a Taylor expansion of $\mathfrak{S} \left(\tilde{q}_n \right)$,

$$\mathfrak{S} \left(\tilde{q}_n \right) = \mathfrak{S} \left(\bar{q} \right) + \mathrm{D}\mathfrak{S} \left(\bar{q} \right) \left[\tilde{q}_n - \bar{q} \right] + \frac{1}{2} \mathrm{D}^2 \mathfrak{S} \left(\nu \right) \left[\tilde{q}_n - \bar{q}, \tilde{q}_n - \bar{q} \right],$$

for some $\nu \in U$, to see that

$$\left(C_s + C_g \right) K_s^n \ge \left| \mathfrak{S} \left(\tilde{q}_n \right) - \mathfrak{S} \left(\bar{q} \right) \right|$$

$$= \left| \mathfrak{S} \left(\bar{q} \right) + \mathrm{D}\mathfrak{S} \left(\bar{q} \right) \left[\tilde{q}_n - \bar{q} \right] + \frac{1}{2} \mathrm{D}^2 \mathfrak{S} \left(\bar{q} \right) \left[\tilde{q}_n - \bar{q}, \tilde{q}_n - \bar{q} \right] - \mathfrak{S} \left(\bar{q} \right) \right|.$$

But

$$\mathrm{D}\mathfrak{S} \left(\bar{q} \right) \left[\tilde{q}_n - \bar{q} \right] = \int_0^h \frac{\partial L}{\partial q} \left(\bar{q}, \dot{\bar{q}} \right) \left(\tilde{q}_n - \bar{q} \right) + \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \left(\dot{\tilde{q}}_n - \dot{\bar{q}} \right) \mathrm{d}t$$

$$= \int_0^h \left(\frac{\partial L}{\partial q} \left(\bar{q}, \dot{\bar{q}} \right) - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \right) \cdot \left(\tilde{q}_n - \bar{q} \right) \mathrm{d}t$$

$$= 0,$$

because $\tilde{q}_n \left(0 \right) = \bar{q} \left(0 \right)$ and $\tilde{q}_n \left(h \right) = \bar{q} \left(h \right)$ by definition (note that this implies $\left(\tilde{q}_n - \bar{q} \right) \in H_0^1 \left(\left[0, h \right], Q \right)$). Then

$$\left(C_s + C_g \right) K_s^n \ge \left| \mathrm{D}^2 \mathfrak{S} \left(\nu \right) \left[\tilde{q}_n - \bar{q}, \tilde{q}_n - \bar{q} \right] \right|$$

$$\ge C_f \left\| \tilde{q}_n - \bar{q} \right\|_{W^{1,1}([0,h])}^2$$

$$C \sqrt{K_s}^n \ge \left\| \tilde{q}_n - \bar{q} \right\|_{W^{1,1}([0,h])}^2$$

where $C = \frac{C_s + C_g}{C_f}$. □

This result shows that Galerkin curves converge to the true solution geometrically with n-refinement, albeit with a larger geometric constant, and hence a slower rate. By simply replacing the bounds (18) and (19) from Theorem 3.4 with those from Theorem 3.3 and the term $C_s K_s^n$ with $C_{op} h^p$, an identical argument shows that Galerkin curves converge at half the optimal rate with h-refinement.

Theorem 3.7. (Convergence of Galerkin Curves with h-Refinement) *Under the same assumptions as Theorem 3.3, if at $\bar{q}$, the action is twice Frechet differentiable, and if the second Frechet derivative of the action $\mathrm{D}^2 \mathcal{S} \left(\cdot \right) \left[\cdot, \cdot \right]$ is coercive with a constant C_f independent of h in a neighborhood U of $\bar{q}$, for all curves $\delta q \in H_0^1 \left(\left[0, h \right], Q \right)$, then if the discrete Lagrange map has error $\mathcal{O} \left(h^{p+1} \right)$, the Galerkin curves have error at most $\mathcal{O} \left(h^{\frac{p+1}{2}} \right)$ in $\left\| \cdot \right\|_{W^{1,1}([0,h])}$. If C_f is a function of h, this bound becomes $\mathcal{O} \left(C_f \left(h \right)^{-1} h^{\frac{p+1}{2}} \right)$.*

Like the requirement that the stationary points of the actions are minimizers, the requirement that the second Frechet derivative of the action is coercive may appear quite strong at first. Again, the coercivity will depend on the properties of the Lagrangian L, but we can establish that for Lagrangians of the canonical form,

$$L(q, \dot{q}) = \frac{1}{2} \dot{q}^T M \dot{q} - V(q),$$

there exists a time step $[0, h]$ over which the action is coercive on $H_0^1([0, h], Q)$.

Theorem 3.8. (Coercivity of the Action) *For Lagrangian of the form*

$$L(q, \dot{q}) = \frac{1}{2} \dot{q}^T M \dot{q} - V(q),$$

where M is symmetric positive-definite, and the second derivatives of $V(q)$ are bounded, there exists an interval $[0, h]$ over which the action is coercive over $H_0^1([0, h], Q)$, that is,

$$D^2 \mathfrak{S}(\nu)[\delta q, \delta q] \geq C_f \|\delta q\|^2_{W^{1,1}([0,h])},$$

for any $\delta q \in H_0^1([0, h], Q)$ and any $\nu \in C^2([0, h], Q)$.

Proof. First, we note that if

$$\mathfrak{S}(\nu) = \int_0^h \frac{1}{2} \dot{\nu}^T M \dot{\nu} - V(\nu),$$

then

$$D^2 \mathfrak{S}(\nu)[\delta q, \delta q] = \int_0^h \delta \dot{q}^T M \delta \dot{q} - \delta q^T H(\nu) \delta q \, \mathrm{d}t$$

$$= \int_0^h \delta \dot{q}^T M \delta \dot{q} \, \mathrm{d}t - \int_0^h \delta q^T H(\nu) \delta q \, \mathrm{d}t$$

where $H(\nu)$ is the Hessian of $V(\nu)$ at the point ν. Since M is symmetric positive-definite, and the second derivatives of $V(\cdot)$ are bounded, then there exists C_r and m such that:

$$\int_0^h \delta \dot{q}^T M \delta \dot{q} \, \mathrm{d}t \geq \int_0^h m \delta \dot{q}^T \delta \dot{q} \, \mathrm{d}t$$

$$(20) \qquad \int_0^h \delta q^T H(\nu) \delta q \, \mathrm{d}t \leq \int_0^h D C_r \delta q^T \delta q \, \mathrm{d}t,$$

(see (17) for a derivation of (20)). Hence,

$$D^2 \mathfrak{S}(\nu)[\delta q, \delta q] \geq \int_0^h m \delta \dot{q}^T \delta \dot{q} \, \mathrm{d}t - \int_0^h D C_r f^T f \, \mathrm{d}t$$

$$(21) \qquad = \frac{1}{2} m \int_0^h \delta \dot{q}^T \delta \dot{q} \, \mathrm{d}t + \frac{1}{2} m \int_0^h \delta \dot{q}^T \delta \dot{q} \, \mathrm{d}t - D C_r \int_0^h \delta q^T \delta q \, \mathrm{d}t.$$

Considering the last two terms in (21), and noting that $\delta q \in H_0^1([0, h], Q)$, we make use of the Poincaré inequality to derive:

$$\frac{1}{2} m \int_0^h \delta \dot{q}^T \delta \dot{q} \, \mathrm{d}t - D C_r \int_0^h \delta q^T \delta q \, \mathrm{d}t \geq \frac{m \pi^2}{2 h^2} \int_0^h \delta q^T \delta q \, \mathrm{d}t - n C_r \int_0^h \delta q^T \delta q \, \mathrm{d}t$$

$$(22) \qquad \geq \left(\frac{m \pi^2}{2 h^2} - D C_r \right) \int_0^h \delta q^T \delta q \, \mathrm{d}t.$$

Thus, substituting (22) in for the last two terms of (21), we conclude:

$$D^2 \mathfrak{S}(q, \dot{q})[\delta q, \delta q] \geq \left(\frac{m \pi^2}{2 h^2} - D C_r \right) \int_0^h \delta q^T \delta q \, \mathrm{d}t + \frac{m}{2} \int_0^h \delta \dot{q}^T \delta \dot{q} \, \mathrm{d}t$$

$$\geq \min \left(\frac{m}{2}, \left(\frac{m \pi^2}{2 h^2} - D C_r \right) \right) \left(\int_0^h \delta q^T \delta q \, \mathrm{d}t + \int_0^h \delta \dot{q}^T \delta \dot{q} \, \mathrm{d}t \right)$$
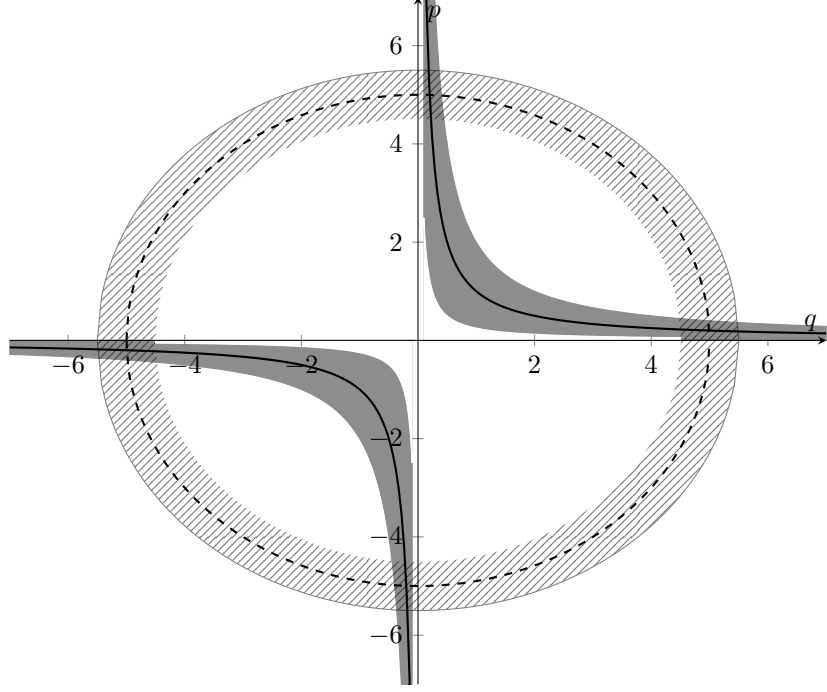
17

FIGURE 2. Conserved and approximately conserved Noether quantities and the resulting constrained solution space. Suppose that both $p^T q = 1$ and $p^2 + q^2 = 5$ were conserved quantities for a certain Lagrangian. Then the solutions of the Euler-Lagrange equations would be constrained to the intersections of these two constant surfaces in phase space; in the above diagram, this is the intersection of the dashed and solid lines. If these quantities were conserved up to a fixed error along a numerical solution, then the numerical solution would be constrained to the intersection of the shaded regions in the above figure. The constraint of the numerical solution to these regions is what leads to the many excellent qualities of variational integrators.

$$= \min\left(\frac{m}{2}, \left(\frac{m\pi^2}{2h^2} - DC_r\right)\right)\left(\|\delta q\|^2_{L^2([0,h])} + \|\delta \dot{q}\|^2_{L^2([0,h])}\right),$$

and making use of Hölder's inequality, we see that $\|\delta q\|_{L^2([0,h])} \geq h^{\frac{1}{2}} \|\delta q\|_{L^1([0,h])}$, thus

$$\begin{aligned}
\mathrm{D}^2\mathfrak{S}(q,\dot{q})[\delta q, \delta q] &\geq \min\left(\frac{m}{2}, \left(\frac{m\pi^2}{2h^2} - DC_r\right)\right)\left(h\|\delta q\|^2_{L^1([0,h])} + h\|\delta \dot{q}\|^2_{L^1([0,h])}\right) \\
&\geq \min\left(\frac{mh}{2}, \left(\frac{m\pi^2}{2h} - hDC_r\right)\right)\frac{1}{2}\left(\|\delta q\|_{L^1([0,h])} + \|\delta \dot{q}\|_{L^1([0,h])}\right)^2 \\
&= \min\left(\frac{mh}{4}, \left(\frac{m\pi^2}{4h} - hDC_r\right)\right)\|\delta q\|^2_{W^{1,1}([0,h])}
\end{aligned}$$

which establishes the coercivity result. $\qquad\qquad\square$

3.4.2. *Noether Quantities.* One of the great advantages of using variational integrators for problems in geometric mechanics is that by construction they have a rich geometric structure which helps lead to excellent long term and qualitative behavior. An important geometric feature of variational integrators is the preservation of discrete Noether quantities, which are invariants that are derived from symmetries of the action. These are analogous to the more familiar Noether quantities of geometric mechanics in the continuous case. We quickly recall Noether's theorem in both the discrete and continuous case, which will also help define the

notation used throughout the proofs that follow. The proofs of both these theorems can be found in Hairer et al. [9].

Theorem 3.9. (Noether's Theorem) *Consider a system with Hamiltonian $H(p, q)$ and Lagrangian $L(q, \dot{q})$. Suppose $\{g_s : s \in \mathbb{R}\}$ is a one-parameter group of transformations which leaves the Lagrangian invariant. Let*

$$a(q) = \left.\frac{d}{ds}\right|_{s=0} g_s(q)$$

be defined as the vector field with flow $g_s(q)$, referred to as the infinitesimal generator, and define the canonical momentum

$$p = \frac{\partial L}{\partial \dot{q}}(q, \dot{q}).$$

Then

$$I(p, q) = p^T a(q)$$

is a first integral of the Hamiltonian system.

Theorem 3.10. (Discrete Noether's Theorem) *Suppose the one-parameter group of transformations leaves the discrete Lagrangian $L_d(q_k, q_{k+1})$ invariant for all (q_k, q_{k+1}). Then:*

$$p_{k+1}^T a(q_{k+1}) = p_k^T a(q_k)$$

where

$$p_k = -D_1 L_d(q_k, q_{k+1}),$$
$$p_{k+1} = D_2 L_d(q_k, q_{k+1}).$$

For the remainder of this section, we will refer to $I(q, p)$ as the *Noether quantity* and $p_n^T a(q_n) = p_{n+1}^T a(q_{n+1})$ as the *discrete Noether quantity*.

For Galerkin variational integrators, it is possible to bound the error of the Noether quantities along the Galerkin curve from the behavior of the analogous discrete Noether quantities of the discrete problem and, more importantly, this bound is independent of the number of time steps that are taken in the numerical integration. This is significant because it offers insight into the excellent behavior of spectral variational integrators even over long periods of integration.

The proof of convergence and near preservation of Noether quantities is broken into three major parts. First, we note that on step k of a numerical integration the discrete Noether quantity arises from a function of the Galerkin curve and the initial point of the one-step map (q_{k-1}, q_k), and that a bound exists for the difference of this discrete Noether quantity evaluated on the Galerkin curve and evaluated on the local exact solution to the Euler-Lagrange equations $\bar{q}$. Second, we show that a bound exists for the difference of the discrete Noether quantity on the local exact solution of the Euler-Lagrange equations and the value of the Noether quantity of the local exact solution, which is conserved along the flow of the Euler-Lagrange vector field. Finally, we show that under certain smoothness conditions, there exists a point-wise bound between the Noether quantity evaluated on the Galerkin curve and the Noether quantity evaluated on the local exact solution. Thus, we establish a point-wise bound between the Noether quantity evaluated on the Galerkin curve and the discrete Noether quantity, and a bound between the discrete Noether quantity and the Noether quantity, which leads to a point-wise bound between the Noether quantity evaluated on the Galerkin curve, and the Noether quantity which is conserved along the global flow of the Euler-Lagrange vector field.

Throughout this section we will make the simplifying assumptions that

$$\tilde{q}_n = \sum_{i=1}^{n} q_k^i \phi_i$$

where $q_k^1 = q_k$, and thus

$$\frac{\partial \tilde{q}_n}{\partial q_k} = \phi_1.$$

This assumption significantly simplifies the analysis.

We begin by bounding the discrete Noether quantity by a function of the local exact solution of the Euler-Lagrange equations.

Lemma 3.3. (Bound on Discrete Noether Quantity) *Define the Galerkin Noether map as:*

$$I_d \left(q\left(t\right), q_k \right) = - \left(h \sum_{j=1}^{n} b_j \left[\frac{\partial L}{\partial q} \left(q, \dot{q} \right) \phi_1 + \frac{\partial L}{\partial \dot{q}} \left(q, \dot{q} \right) \dot{\phi}_1 \right] \right)^T a \left(q_k \right)$$

and note that the discrete Noether quantity is given by

$$I_d \left(\tilde{q}_n, q_k \right) = p_n^T a \left(q_k \right).$$

Assuming the quadrature accuracy of Theorem (3.4) with n-refinement and Theorem (3.3) with h-refinement, if $\frac{\partial L}{\partial q} \left(q, \dot{q} \right)$, $\frac{\partial L}{\partial \dot{q}} \left(q, \dot{q} \right)$ and $\frac{d}{dt} \frac{\partial L}{\partial \dot{q}}$ are Lipschitz continuous, $\| \phi_1 \|_{L^\infty([0,h])}$ is bounded with n refinement, and $\| \tilde{q}_n - \bar{q} \|_{W^{1,1}([0,h])}$ is bounded below by the quadrature error, then

$$\left| I_d \left(\tilde{q}_n, q_k \right) - I_d \left(\bar{q}, q_k \right) \right| \leq C \left| a \left(q_k \right) \right| \left(\| \tilde{q}_n - \bar{q} \|_{W^{1,1}([0,h])} + \| \tilde{q}_n - \bar{q} \|_{L^\infty([0,h])} + \left\| \dot{\tilde{q}}_n - \dot{\bar{q}} \right\|_{L^\infty([0,h])} \right)$$

for some C independent of n and h.

Proof. We begin by expanding the definitions of the discrete Noether quantity:

$$\left| I_d \left(\tilde{q}_n, q_k \right) - I_d \left(\bar{q}, q_k \right) \right| = \left| h \left(\sum_{j=1}^{m} b_j \left[\frac{\partial L}{\partial q} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) \phi_1 + \frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) \dot{\phi}_1 \right] \right)^T a \left(q_k \right) \right.$$

$$\left. - \left(h \sum_{j=1}^{m} b_j \left[\frac{\partial L}{\partial q} \left(\bar{q}, \dot{\bar{q}} \right) \phi_1 + \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \dot{\phi}_1 \right] \right)^T a \left(q_k \right) \right|$$

$$= \left| \left(h \sum b_j \left[\left(\frac{\partial L}{\partial q} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial q} \left(\bar{q}, \dot{\bar{q}} \right) \right) \phi_1 - \left(\frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \right) \dot{\phi}_1 \right] \right)^T a \left(q_k \right) \right|$$

$$\leq \left| h \sum_{j=1}^{m} b_j \left[\left(\frac{\partial L}{\partial q} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial q} \left(\bar{q}, \dot{\bar{q}} \right) \right) \phi_1 - \left(\frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \right) \dot{\phi}_1 \right] \right| \left| a \left(q_k \right) \right|.$$

Now we introduce the function $e_q \left(\cdot, \cdot \right)$ which gives the error of the quadrature rule, and thus

$$\left| I_d \left(\tilde{q}_n, q_k \right) - I_d \left(\bar{q}, q_k \right) \right| \leq \left| \int_0^h \left(\frac{\partial L}{\partial q} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial q} \left(\bar{q}, \dot{\bar{q}} \right) \right) \phi_1 - \left(\frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \right) \dot{\phi}_1 \mathrm{dt} \right.$$

$$\left. + e_q \left(\tilde{q}_n - \bar{q}, \dot{\tilde{q}}_n - \dot{\bar{q}} \right) \right| \left| a \left(q_k \right) \right|.$$

Integrating by parts, we get:

$$\left| I_d \left(\tilde{q}_n, q_k \right) - I_d \left(\bar{q}, q_k \right) \right| \leq \left| \int_0^h \left(\frac{\partial L}{\partial q} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial q} \left(\bar{q}, \dot{\bar{q}} \right) \right) \phi_1 - \frac{\mathrm{d}}{\mathrm{dt}} \left(\frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \right) \phi_1 \mathrm{dt} \right.$$

$$\left. + \left(\frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \frac{\partial L}{\partial \dot{q}} \left(\bar{q}, \dot{\bar{q}} \right) \right) \phi_1 \Big|_0^h + e_q \left(\tilde{q}_n - \bar{q}, \dot{\tilde{q}}_n - \dot{\bar{q}} \right) \right| \left| a \left(q_k \right) \right|.$$

Introducing the Lipschitz constants L_1 for $\frac{\partial L}{\partial q}$, L_2 for $\frac{\partial L}{\partial \dot{q}}$, and L_3 for $\frac{d}{dt} \frac{\partial L}{\partial \dot{q}}$,

$$\left| I_d \left(\tilde{q}_n, q_k \right) - I_d \left(\bar{q}, q_k \right) \right| \leq \left(\int_0^h \left(L_1 + L_3 \right) \left| \left(\tilde{q}_n, \dot{\tilde{q}}_n \right) - \left(\bar{q}, \dot{\bar{q}} \right) \right| \left| \phi_1 \right| \mathrm{dt} + 2 L_2 \left(\| \phi_1 \|_{L^\infty([0,h])} \right) \right.$$

$$\left. \left(\| \tilde{q}_n - \bar{q} \|_{L^\infty([0,h])} + \left\| \dot{\tilde{q}}_n - \dot{\bar{q}} \right\|_{L^\infty([0,h])} \right) + e_q \left(\tilde{q}_n - \bar{q}, \dot{\tilde{q}}_n - \dot{\bar{q}} \right) \right) \left| a \left(q_k \right) \right|$$

$$\leq (L_1 + L_3) \|\phi_1\|_{L^\infty([0,h])} |a(q_k)| \left(\int_0^h \left| (\tilde{q}_n, \dot{\tilde{q}}_n) - (\bar{q}, \dot{\bar{q}}) \right| dt \right)$$

$$+ 2L_2 \left(\|\phi_1\|_{L^\infty([0,h])} \right) |a(q_k)| \left(\|\tilde{q}_n - \bar{q}\|_{L^\infty([0,h])} + \|\dot{\tilde{q}}_n - \dot{\bar{q}}\|_{L^\infty([0,h])} \right)$$

$$+ e_q \left(\tilde{q}_n - \bar{q}, \dot{\tilde{q}}_n - \dot{\bar{q}} \right) |a(q_k)|.$$

We now make the simplification that the quadrature error $|e_q(\cdot, \cdot)|$ serves as a lower bound for $\|\tilde{q}_n - \bar{q}\|_{W^{1,1}([0,h])}$. While this may not strictly hold, all of our estimates on the convergence for $\tilde{q}_n$ imply this bound, and hence it is a reasonable simplification for establishing convergence in this case. Now, note that $\|\phi_1\|_{L^\infty([0,h])}$ is invariant under h rescaling, and let

$$C = \max(L_1 + L_3, 2L_2) \|\phi_1\|_{L^\infty([0,h])} + 1$$

to get

$$|I_d(\tilde{q}_n, q_k) - I_d(\bar{q}, q_k)| \leq C |a(q_k)| \left(\|\tilde{q}_n - \bar{q}\|_{W^{1,1}([0,h])} + \|\tilde{q}_n - \bar{q}\|_{L^\infty([0,h])} + \|\dot{\tilde{q}}_n - \dot{\bar{q}}\|_{L^\infty([0,h])} \right)$$

which establishes the result. $\qquad\square$

Lemma 3.3 establishes a bound between the discrete Noether quantity and $I_d(\bar{q}, q_k)$. The next step is to establish a bound between $I_d(\bar{q}, q_k)$ and the Noether quantity.

Lemma 3.4. (Error Between Discrete Noether Quantity and True Noether Quantity) *Assume that $\phi_1(0) = 1$ and $\phi_1(h) = 0$, and that the sequence $\{|a(q_k)|\}_{k=1}^N$ is bounded by a constant C_a which is independent of N. Let*

$$\bar{p}(t) = \frac{\partial L}{\partial \dot{q}} (\bar{q}(t), \dot{\bar{q}}(t)).$$

Once again, let the error of the quadrature rule be given by $e_q(\cdot, \cdot)$. Then

$$\left| I^d(\bar{q}, q_k) - I(\bar{p}(t), \bar{q}(t)) \right| \leq C_a |e_q(\bar{q}, \dot{\bar{q}})|$$

for any $t \in [0, h]$.

Proof. First, we note that since $\bar{q}$ solves the Euler-Lagrange equations exactly, $I(\bar{p}(t), \bar{q}(t))$ is a conserved quantity along the flow, so it suffices to show the inequality holds for $t = 0$. We begin by expanding:

$$\left| I^d(\bar{q}, q_k) - I(\bar{p}(0), \bar{q}(0)) \right| = \left| -h \left(\sum_{j=1}^m b_j \frac{\partial L}{\partial q}(\bar{q}, \dot{\bar{q}}) \phi_1 + \frac{\partial L}{\partial \dot{q}}(\bar{q}, \dot{\bar{q}}) \dot{\phi}_1 \right)^T a(q_k) - \bar{p}(0)^T a(\bar{q}(0)) \right|$$

$$= \left| -\left(\int_0^h \frac{\partial L}{\partial q}(\bar{q}, \dot{\bar{q}}) \phi_1 + \frac{\partial L}{\partial \dot{q}}(\bar{q}, \dot{\bar{q}}) \dot{\phi}_1 dt + e_q(\bar{q}, \dot{\bar{q}}) \right)^T a(q_k) - \bar{p}(0)^T a(\bar{q}(0)) \right|$$

$$= \left| -\left(\int_0^h \left(\frac{\partial L}{\partial q}(\bar{q}, \dot{\bar{q}}) - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}(\bar{q}, \dot{\bar{q}}) \right) \phi_1 dt + \frac{\partial L}{\partial \dot{q}}(\bar{q}(h), \dot{\bar{q}}(h)) \phi_1(h) \right.\right.$$

$$\left.\left. - \frac{\partial L}{\partial \dot{q}}(\bar{q}(0), \dot{\bar{q}}(0)) \phi_1(0) + e_q(\bar{q}, \dot{\bar{q}}) \right)^T a(q_k) - \bar{p}(0)^T a(\bar{q}(0)) \right|$$

Since $\bar{q}(t)$ solves the Euler-Lagrange equations, $\phi_1(0) = 1$ and $\phi_1(h) = 0$, and $\bar{q}(0) = q_k$,

$$\left| I^d(\bar{q}, q_k) - I(\bar{p}(0), \bar{q}(0)) \right| = \left| \left(\frac{\partial L}{\partial \dot{q}}(\bar{q}(0), \dot{\bar{q}}(0)) \right)^T a(q_k) + (e_q(\bar{q}, \dot{\bar{q}}))^T a(q_k) - \bar{p}(0)^T a(q_k) \right|$$

$$= \left| (\bar{p}(0))^T a(q_k) + (e_q(\bar{q}, \dot{\bar{q}}))^T a(q_k) - (\bar{p}(0))^T a(q_k) \right|$$

$$= \left| e_q(\bar{q}, \dot{\bar{q}})^T a(q_k) \right|$$

$$\leq |e_q(\bar{q}, \dot{\bar{q}})| |a(q_k)|$$

21

$$\leq C_a \left| e_q\left(\bar{q}, \dot{\bar{q}} \right) \right|$$

which yields the desired bound. $\qquad\square$

Once again, if we assume that the quadrature error serves as a lower bound for the Sobolev error, combining the bounds from (3.3) and (3.4) yields:

$$\left| I_d\left(\tilde{q}_n, q_k \right) - I\left(\bar{p}\left(t \right), \bar{q}\left(t \right) \right) \right| \leq 2CC_a \left(\left\| \tilde{q}_n - \bar{q} \right\|_{W^{1,1}([0,h])} + \left\| \tilde{q}_n - \bar{q} \right\|_{L^\infty([0,h])} + \left\| \dot{\tilde{q}}_n - \dot{\bar{q}} \right\|_{L^\infty([0,h])} \right).$$

This bound serves two purposes; the first is to establish a bound between the discrete Noether quantity and the Noether quantity computed on the local exact solution $\bar{q}$. The second is to establish a bound between the discrete Noether quantity after one step and the Noether quantity computed on the initial data:

$$\left| I_d\left(\tilde{q}_n, q_1 \right) - I\left(p\left(0 \right), q\left(0 \right) \right) \right| \leq 2CC_a \left(\left\| \tilde{q}_n - \bar{q} \right\|_{W^{1,1}([0,h])} + \left\| \tilde{q}_n - \bar{q} \right\|_{L^\infty([0,h])} + \left\| \dot{\tilde{q}}_n - \dot{\bar{q}} \right\|_{L^\infty([0,h])} \right),$$

since for (q_1, q_2), $\bar{q}$ is the *global* exact flow of the Euler-Lagrange equations.

The difference between these two bounds is subtle but important; by establishing a bound between the discrete Noether quantity and the Noether quantity associated with the initial conditions, on any step of the method we can bound the error between the discrete Noether quantity and the Noether quantity associated with the *global* exact flow. By establishing the bound between the discrete Noether quantity and the Noether quantity associated with $\bar{q}$ at any step, we can bound the error between the Noether quantity associated with the local exact flow $\bar{q}$ and the true Noether quantity conserved along the global exact flow:

$$\left| I\left(\bar{p}\left(t \right), \bar{q}\left(t \right) \right) - I\left(p\left(0 \right), q\left(0 \right) \right) \right| \leq \left| I\left(\bar{p}\left(t \right), \bar{q}\left(t \right) \right) - I_d\left(\tilde{q}_n, q_k \right) \right| + \left| I_d\left(\tilde{q}_n, q_k \right) - I\left(p\left(0 \right), q\left(0 \right) \right) \right|$$

$$(23) \qquad\qquad \leq 4CC_a \left(\left\| \tilde{q}_n - \bar{q} \right\|_{W^{1,1}([0,h])} + \left\| \tilde{q}_n - \bar{q} \right\|_{L^\infty([0,h])} + \left\| \dot{\tilde{q}}_n - \dot{\bar{q}} \right\|_{L^\infty([0,h])} \right)$$

for any $t_0 \in [0, h]$ on any time step k. Because the local exact flow $\bar{q}$ is generated from boundary conditions (q_k, q_{k+1}) which only approximate the boundary conditions of the true flow, there is no guarantee that the Noether quantity associated with $\bar{q}$ will be the same step to step, only that it will be conserved within each time step. However, because there is a bound between the Noether quantity associated with $\bar{q}$ and the discrete Noether quantity at every time step, the discrete Noether quantity and the Noether quantity associated with the exact flow, and because the Noether quantity is conserved point-wise along $\bar{q}$ on each time step, there exists a bound between the Noether quantity associated with each point of the local exact flow and the Noether quantity associated with the true solution.

We finally arrive at the desired result, which is a theorem that bounds the error between the Noether quantity along the Galerkin curve and the true Noether quantity. It is significant because not only does it bound the error of the Noether quantity, but the bound is independent of the number of steps taken, and hence will not grow even for extremely long numerical integrations.

Theorem 3.11. (Convergence of Conserved Noether Quantities) *Define*

$$\tilde{p}_n = \frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n, \dot{\tilde{q}}_n \right).$$

Under the assumptions of Lemmas (3.3 - 3.4), if the Noether map $I\left(p, q \right)$ is Lipschitz continuous in both its arguments, then there exists a constant C_v independent N, the number of method steps, such that:

$$\left| I\left(p\left(0 \right), q\left(0 \right) \right) - I\left(\tilde{p}_n\left(t \right), \tilde{q}_n\left(t \right) \right) \right| \leq C_v \left(\left\| \tilde{q}_n - \bar{q} \right\|_{W^{1,1}([0,h])} + \left\| \tilde{q}_n - \bar{q} \right\|_{L^\infty([0,h])} + \left\| \dot{\tilde{q}}_n - \dot{\bar{q}} \right\|_{L^\infty([0,h])} \right).$$

for any $t \in [0, Nh]$.

Proof. We begin by introducing the Noether quantity evaluated at t on the local exact flow, $\bar{q}$:

$$(24) \qquad \left| I\left(p\left(0 \right), q\left(0 \right) \right) - I\left(\tilde{p}_n\left(t \right), \tilde{q}_n\left(t \right) \right) \right| \leq \left| I\left(\tilde{p}_n\left(t \right), \tilde{q}_n\left(t \right) \right) - I\left(\bar{p}\left(t \right), \bar{q}\left(t \right) \right) \right|$$
$$+ \left| I\left(\bar{p}\left(t \right), \bar{q}\left(t \right) \right) - I\left(p\left(0 \right), q\left(0 \right) \right) \right|.$$

Considering the first term in (24), let L_4 be the Lipschitz constant for $I\left(\cdot, \cdot \right)$. Then

$$\left| I\left(\tilde{p}_n\left(t \right), \tilde{q}_n\left(t \right) \right) - I\left(\bar{p}\left(t \right), \bar{q}\left(t \right) \right) \right| \leq L_4 \left| \left(\tilde{p}_n\left(t \right), \tilde{q}_n\left(t \right) \right) - \left(\bar{p}\left(t \right), \bar{q}\left(t \right) \right) \right|$$
$$\leq L_4 \left(\left| \tilde{p}_n\left(t \right) - \bar{p}\left(t \right) \right| + \left| \tilde{q}_n\left(t \right) - \bar{p}\left(t \right) \right| \right)$$

22

$$= L_4 \left(\left| \frac{\partial L}{\partial \dot{q}} \left(\tilde{q}_n \left(t \right), \dot{\tilde{q}}_n \left(t \right) \right) - \frac{\partial L}{\partial \dot{q}} \left(\bar{q} \left(t \right), \dot{\bar{q}}(t) \right) \right| + \left| \tilde{q}_n \left(t \right) - \bar{q} \left(t \right) \right| \right)$$

$$\leq L_4 \left(L_2 \left| \dot{\tilde{q}}_n \left(t \right) - \dot{\bar{q}} \left(t \right) \right| + \left(L_2 + 1 \right) \left| \tilde{q}_n \left(t \right) - \bar{q} \left(t \right) \right| \right)$$

(25)
$$\leq L_4 \left(L_2 + 1 \right) \left(\left\| \tilde{q}_n - \bar{q} \right\|_{L^\infty([0,h])} + \left\| \dot{\tilde{q}}_n - \dot{\bar{q}} \right\|_{L^\infty([0,h])} \right).$$

The second term in (24) is exactly the bound given by (23) and thus combining (25) and (23) in (24) and defining $C_v = 4CC_a + L\left(L_2 + 1 \right)$, we have:

$$\left| I \left(\tilde{p}_n \left(t \right), \tilde{q}_n \left(t \right) \right) - I \left(\bar{p} \left(t \right), \bar{q} \left(t \right) \right) \right| \leq C_v \left(\left\| \tilde{q}_n - \bar{q} \right\|_{W^{1,1}([0,h])} + \left\| \tilde{q}_n - \bar{q} \right\|_{L^\infty([0,h])} + \left\| \dot{\tilde{q}}_n - \dot{\bar{q}} \right\|_{L^\infty([0,h])} \right)$$

which completes the result. $\qquad\square$

The convergence and bounds of the Noether quantity evaluated on the Galerkin curve to that of the true solution is hampered by one issue. While Theorems (3.6) and (3.7) provide estimates for convergence in the Sobolev norm $\left\| \cdot \right\|_{W^{1,1}([0,h])}$, Theorem (3.11) requires estimates in the L^∞ norm. We can establish a bound for $\left\| \tilde{q}_n \left(t \right) - \bar{q} \left(t \right) \right\|_{L^\infty([0,h])}$, but it is much more difficult to establish a general estimate for $\left\| \dot{\tilde{q}}_n \left(t \right) - \dot{\bar{q}} \left(t \right) \right\|_{L^\infty([0,h])}$.

Lemma 3.5. (Bound on L^∞ Norm from Sobolev Norm) *For any $t \in [0,h]$, the following bound holds:*

$$\left| q \left(t \right) \right| \leq \max \left(\frac{1}{h}, 1 \right) \left\| q \right\|_{W^{1,1}([0,h])}$$

and thus

$$\left\| q \right\|_{L^\infty([0,h])} \leq \max \left(\frac{1}{h}, 1 \right) \left\| q \right\|_{W^{1,1}([0,h])}.$$

Proof. This is a basic extension of the arguments from Lemma A.1. in Larsson and Thomée [11], generalizing the lemma from the interval $[0,1]$ to an interval of arbitrary length, $[0,h]$. We note that for any $t, s \in [0,h]$, $q \left(t \right) = q \left(s \right) + \int_s^t \dot{q} \left(u \right) \mathrm{d}u$. Thus:

$$\left| q \left(t \right) \right| \leq \left| q \left(s \right) \right| + \int_0^h \left| \dot{q} \left(u \right) \right| \mathrm{d}u$$
$$\leq \left| q \left(s \right) \right| + \left\| \dot{q} \right\|_{L^1([0,h])}.$$

Now, we integrate with respect to s:

$$\int_0^h \left| q \left(t \right) \right| \mathrm{d}s \leq \int_0^h \left| q \left(s \right) \right| \mathrm{d}s + \int_0^h \left\| \dot{q} \right\|_{L^1([0,h])} \mathrm{d}s$$
$$h \left| q \left(t \right) \right| \leq \left(\left\| q \right\|_{L^1([0,h])} + h \left\| \dot{q} \right\|_{L^1([0,h])} \right).$$

which yields the desired result. $\qquad\square$

Under certain assumptions about the behavior of $\dot{\tilde{q}}_n - \dot{\bar{q}}$, it is possible to establish bounds on the point-wise error of $\dot{\tilde{q}}_n$ from the Sobolev error $\left\| \tilde{q}_n - \bar{q} \right\|_{W^{1,1}([0,h])}$. For example, if the length of time that the error is within a given fraction of the max error is proportional to the length of the interval $[0,h]$, i.e. there exists C_1, C_2 independent of h: i.e.,

$$m \left(\left\{ t \, \middle| \, \left\| \left(\dot{\tilde{q}}_n \left(t \right) - \dot{\bar{q}} \left(t \right) \right) \right\| \geq C_1 \left\| \dot{\tilde{q}}_n \left(t \right) - \dot{\bar{q}} \left(t \right) \right\|_\infty \right\} \right) \geq C_2 h,$$

where m is the Lebesque measure, then it can easily be seen that:

$$\left\| \tilde{q}_n - \bar{q} \right\|_{W^{1,1}([0,h])} \geq \int_0^h \left\| \dot{\tilde{q}}_n \left(t \right) - \dot{\bar{q}} \left(t \right) \right\| \mathrm{d}t \geq C_1 C_2 h \left\| \dot{\tilde{q}}_n - \dot{\bar{q}} \right\|_{L^\infty([0,h])}.$$

While we will not establish here that the $\dot{\tilde{q}}_n$ converges in the L^∞ norm with the same rate that the Galerkin curve converges in the Sobolev norm, our numerical experiments will show that the Noether quantities tend to converge at the same rate as the Galerkin curve.

23

4. Numerical Experiments

To support the results in this paper, as well as to investigate the efficiency and stability of spectral variational integrators, several numerical experiments were conducted by applying spectral variational techniques to well-known variational problems. For each problem, the spectral variational integrator was constructed using a Lagrange interpolation polynomials at n Chebyshev points with a Gauss quadrature rule at $2n$ points. Convergence of both the one-step map and the Galerkin curves was measured using the ℓ^∞ and L^∞ norms respectively, although we record them on the same axis using labeled L^∞ error in a slight abuse of notation. The experiments strongly support the results of this paper, and suggest topics for further investigation.

4.1. Harmonic Oscillator.
The first and simplest numerical experiment conducted was the harmonic oscillator. Starting from the Lagrangian,

$$L(q, \dot{q}) = \frac{1}{2}\dot{q}^2 - \frac{1}{2}q^2,$$

where $q \in \mathbb{R}$, the induced spectral variational discrete Euler-Lagrange equations are linear. Choosing the large time step $h = 20$ over 100 steps yields the expected geometric convergence, and attains very high accuracy, as can be seen in Figure 3. In addition, the max error of the energy converges geometrically, see Figure 4, and does not grow over the time of integration, see Figure 5.

4.2. N-body Problems.
We now turn our attention towards Kepler N-body problems, which are both good benchmark problems and are interesting in their own right. The general form of the Lagrangian for these problems is

$$L(q, \dot{q}) = \frac{1}{2}\sum_{i=1}^{N}\dot{q}_i^T M \dot{q}_i + G\sum_{i=1}^{N}\sum_{j=0}^{i-1}\frac{m_i m_j}{\|q_i - q_j\|},$$

where $q_i \in \mathbb{R}^D$ is the center of mass for body i, G is a gravitational constant, and m_i is a mass constant associated with the body described by q_i.

4.2.1. 2-Body Problem.
The first experiment we will examine is the choice of parameters $D = 2$, $m_1 = m_2 = 1$. Centering the coordinate system at q_1, we choose $q_2(0) = (0.4, 0)$, $\dot{q}_2(0) = (0, 2)$, which has a known closed form solution which is a stable closed elliptical orbit with eccentricity 0.6. Knowing the closed form solution allows us to examine the rate of convergence to the true solution, and when solved with the large time step $h = 2.0$, over 100 steps, the error of the one-step map is $\mathcal{O}(0.56^n)$ with n-refinement and $\mathcal{O}\left(h^{2\lceil\frac{n}{2}\rceil}\right)$ with h-refinement, as can be seen in Figure 6 and Figure 9, respectively. The numerical evidence suggests that our bound for the one-step map with h-refinement is not sharp, as the convergence of the one-step map is always even. Interestingly, it is also possible to observe the different convergence rates of the one-step map and the Galerkin curves with n-refinement, as eventually the Galerkin curves have error approximately $\mathcal{O}(0.74^n)$ while the one-step map has error approximately $\mathcal{O}(0.56^n)$, and $\sqrt{0.56} \approx 0.7483$. However, it appears that the error from the one step map dominates until very high choices of n, and thus it is difficult to observe the error of the Galerkin curves directly with h-refinement, round off error becomes a problem before the error of the Galerkin curves does for smaller choices of n.

The N-body Lagrangian is invariant under the action of $\mathrm{SO}(D)$, which yields the conserved Noether quantity of angular momentum. For the two body problem this is:

$$L(q, \dot{q}) = q_x\dot{q}_y - q_y\dot{q}_x$$

where $q = (q_x, q_y)$. Numerical experiments show that the error of the angular momentum does not grow with the number of steps taken in the integration, Figure 13, but that the error is of the same order as the error of Galerkin curve with n-refinement in Figure 8. With h-refinement, the angular momentum appears to have error $\mathcal{O}\left(h^{\frac{n}{2}} + 2\right)$ in Figure 11. This is interesting because the theoretical bound on the error of the Galerkin curves is $\mathcal{O}\left(h^{\frac{n}{2}}\right)$, and the error of the Noether quantities is theoretically a factor $C(h)$ times the error of the Galerkin curves, where C is the factor that arises in the proof of the convergence of the conserved Noether quantities. Numerical experiments suggest C is $\mathcal{O}(h^2)$ for this problem, but that the Galerkin curves do converge at a rate of $\mathcal{O}\left(h^{\frac{n}{2}}\right)$, which is consistent with of the Galerkin curve error estimate. While this
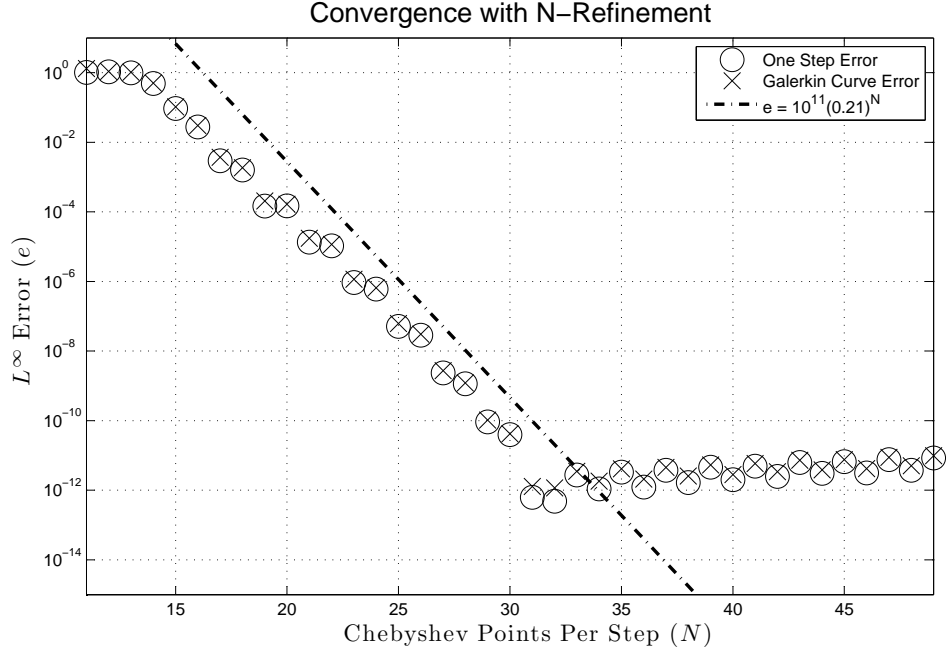
FIGURE 3. Geometric convergence of the spectral variational integration of the harmonic oscillator problem, for 100 steps at step size $h = 20.0$.
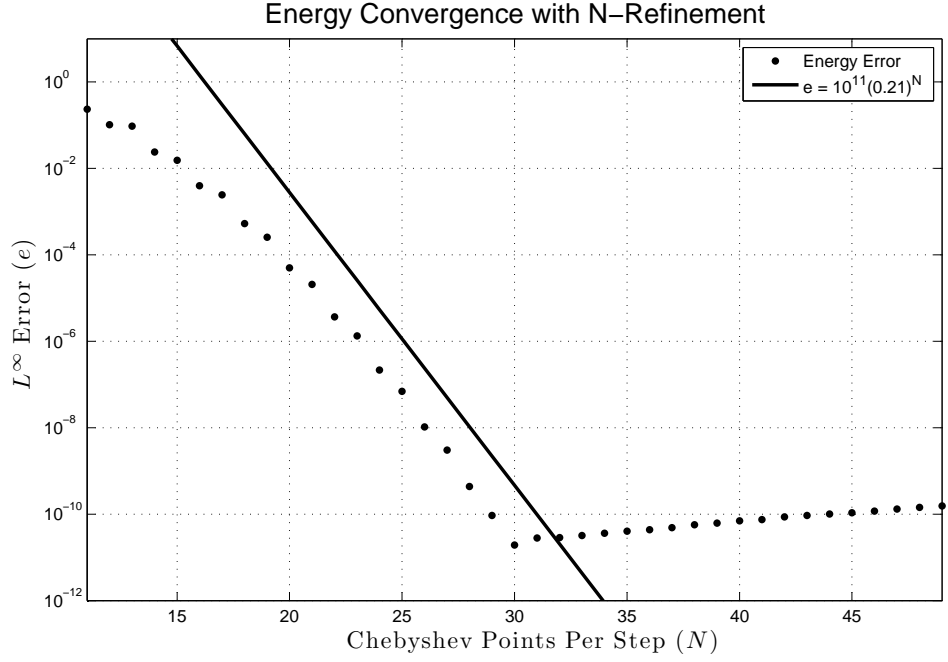

FIGURE 4. Geometric convergence of the energy error of the spectral variational integration of the harmonic oscillator problem for 100 steps at step size $h = 20.0$.
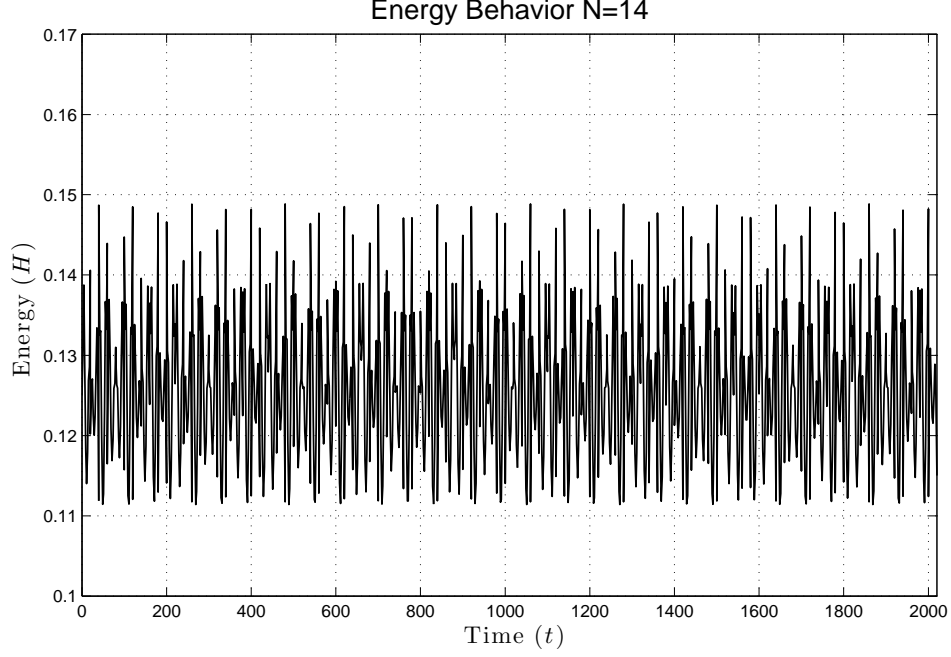
FIGURE 5. Energy Stability of the spectral variational integration of the harmonics oscillator problem. This energy was computed for the integration using $n = 14$ for step size $h = 20.0$.
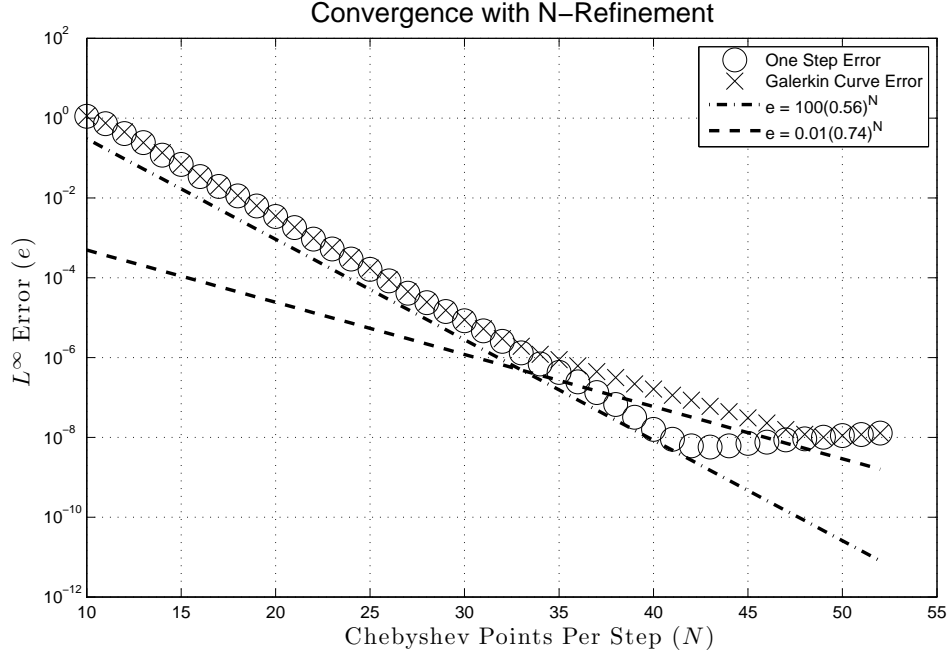
FIGURE 6. Geometric convergence of the Kepler 2-body problem with eccentricity 0.6 over 100 steps of $h = 2.0$. Note that around $n = 32$, the error for the Galerkin curves becomes $\mathcal{O}\left(0.74^n\right)$, while the error for the one-step map is always $\mathcal{O}\left(0.56^n\right)$.
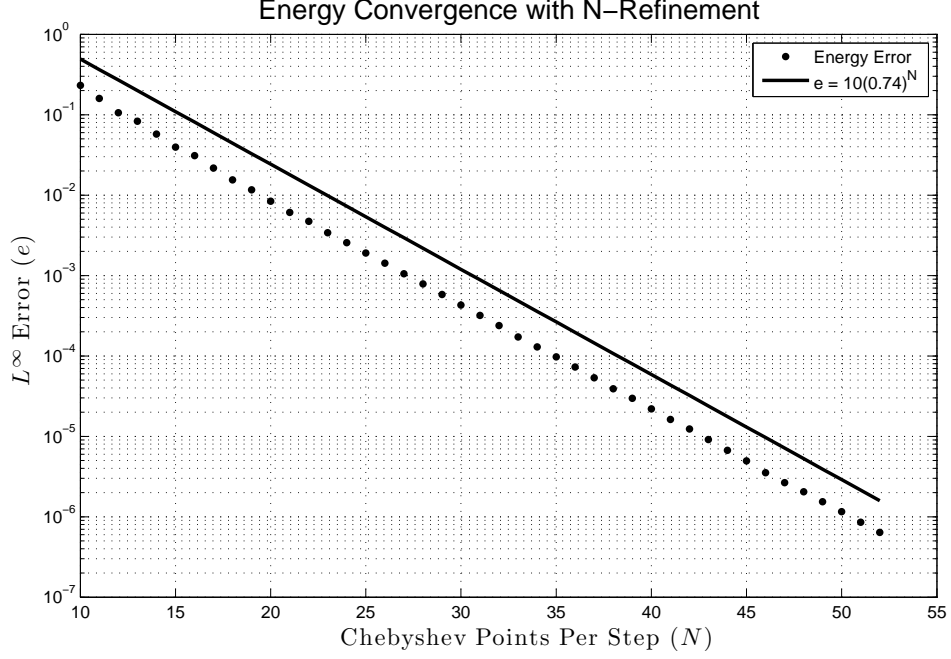
FIGURE 7. Geometric convergence of the Energy Error of Kepler 2-body problem with eccentricity 0.6 over 100 steps of $h = 2.0$. Note that the error is $\mathcal{O}(0.74^n)$, the same as it was for the Galerkin curves.
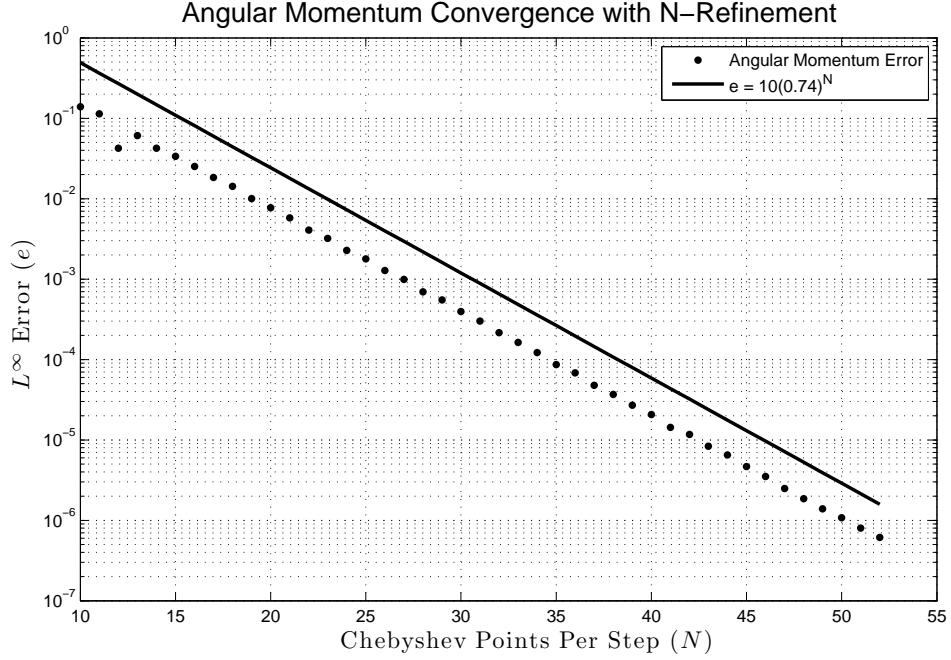
FIGURE 8. Geometric convergence of the angular momentum of the Kepler 2-body problem with eccentricity 0.6 over 100 steps of $h = 2.0$. Again, the error is of the same order as it was the Galerkin curves.

One Step Map Convergence with h−Refinement
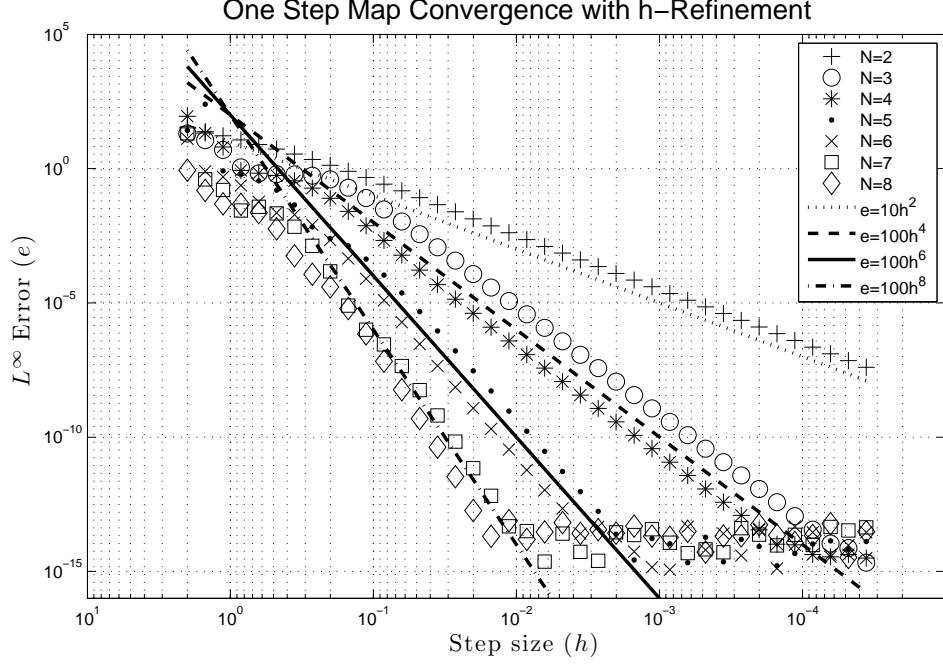
FIGURE 9. Order Optimal convergence of the Kepler 2-body problem with eccentricity 0.6 over 10 steps with h refinement. Note our bound is not sharp, as the error is $\mathcal{O}\left(h^{2\left\lceil \frac{n}{2}\right\rceil}\right)$, where $\lceil \cdot \rceil$ is the ceiling function.

Energy Convergence with h−Refinement
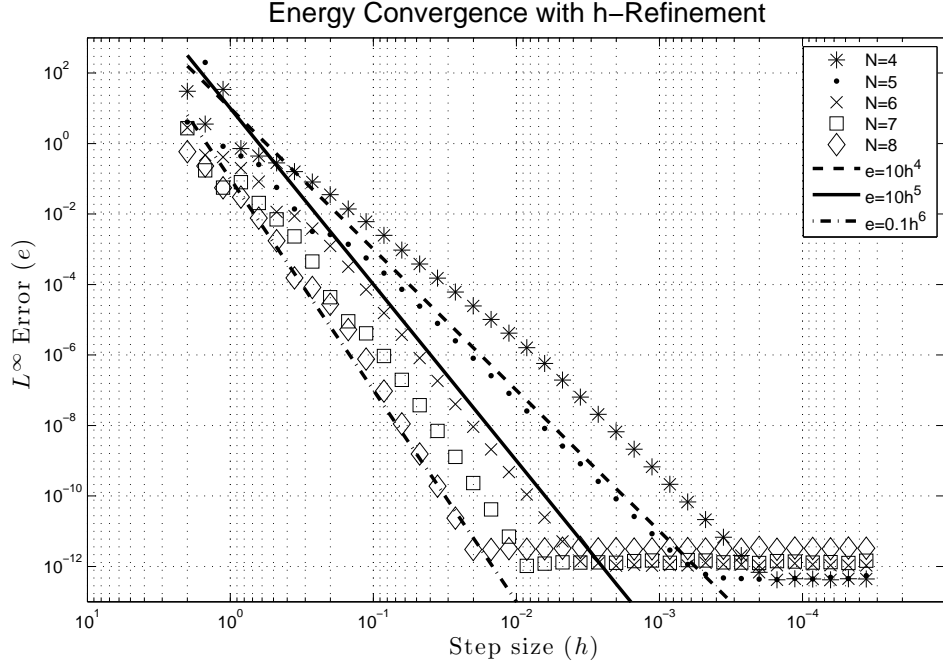
FIGURE 10. Convergence of the Kepler 2-body problem energy with eccentricity 0.6 over 10 steps with h refinement.
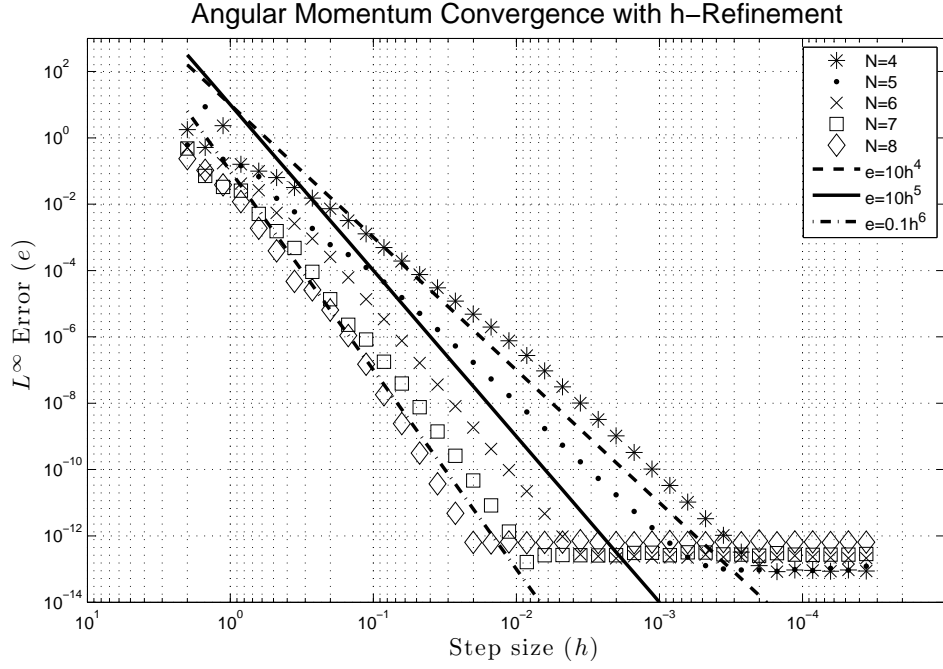
FIGURE 11. Convergence of the angular momentum Kepler 2-body problem with eccentricity 0.6 over 100 steps of $h = 2.0$.
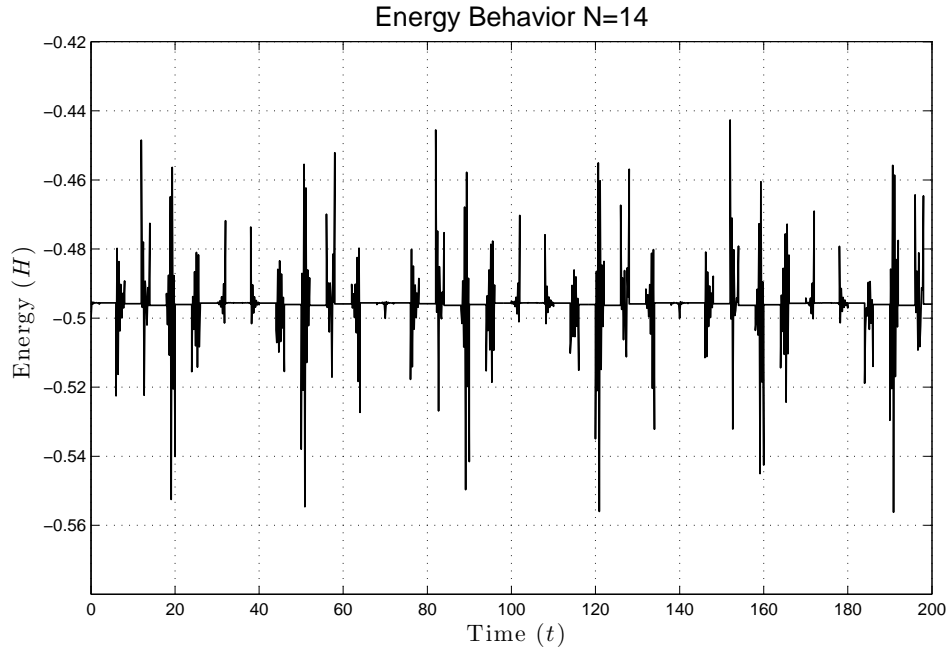
FIGURE 12. Stability of energy for Kepler 2-body problem.
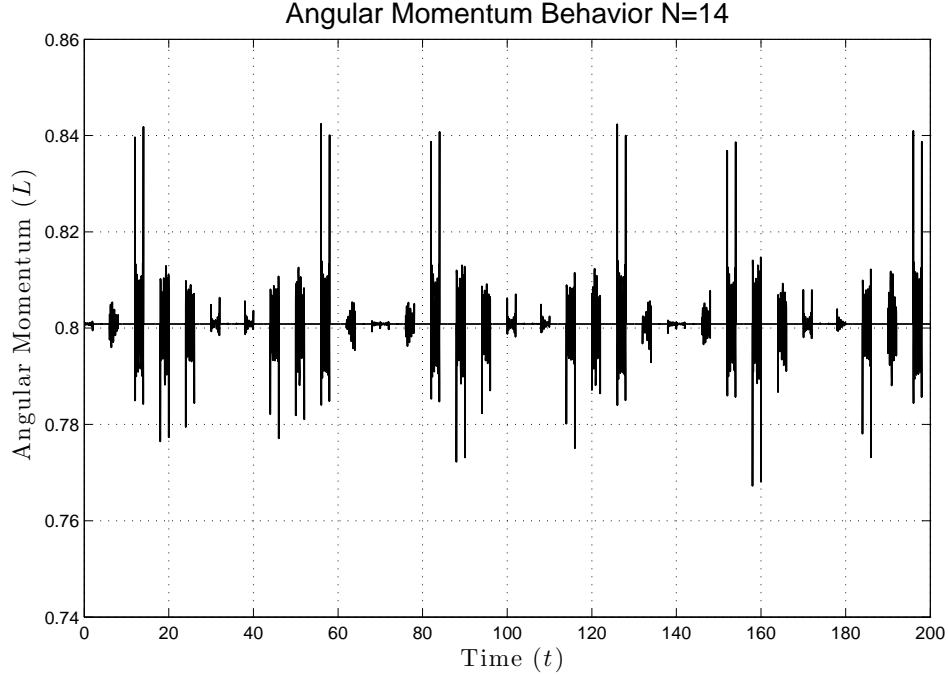
FIGURE 13. Stability of angular momentum for Kepler 2-body problem.

evidence is not conclusive, it is suggestive that the error analysis provides a plausible bound. A careful analysis of the factor C would be an interesting direction for further investigation.

4.2.2. *The Solar System.* To illustrate the excellent stability proprieties of spectral variational integrators, we let $D = 3$, $N = 10$, and use the velocities, positions, and masses of the sun, 8 planet, and the dwarf planet Pluto on January 1, 2000 (as provided by the JPL Solar System ephemeris [23]) as initial configuration parameters for the Kepler system. Taking 100 time steps of $h = 100$ days, the $N = 25$ spectral variational integrator produces a highly stable flow in Figure 14. It should be noted that orbits are closed, stable, and exhibit almost none of the "precession" effects that are characteristic of symplectic integrators, even though the time step is larger than the orbital period of Mercury. Additionally, considering just the outer solar system (Jupiter, Saturn, Uranus, Neptune, Pluto), and aggregating the inner solar system (Sun, Mercury, Venus, Earth, Mars) to a point mass, an $N = 25$ spectral variational integrator taking 100 time steps $h = 1825$ days (5 year steps) produces the orbital flow seen in Figure 15. Again, these are highly stable, precession free orbits. As can be clearly seen, the spectral variational integrator produces extremely stable flows, even for very large time steps.

5. Conclusions and Future Work

In this paper a new numerical method for variational problems was introduced, specifically a symplectic momentum-preserving integrator that exhibits geometric convergence to the true flow of a system under the appropriate conditions. These integrators were constructed under the general framework of Galerkin variational integrators, and made use of the global function paradigm common to many different spectral methods.

Additionally, a general convergence theorem was established for Galerkin type variational integrators, establishing the important result that, under suitable hypotheses, Galerkin variational integrators will inherit the optimal order of convergence permitted by the underlying approximation space used in their construction. This result provides a powerful tool for both constructing and analyzing variational integrators, it provides a methodology for constructing methods of very high order of accuracy, and it also establishes order of convergence for methods that can be viewed as Galerkin variational integrators. It was shown that from the one step map, a continuous approximation to the solution of the Euler-Lagrange equations can be easily
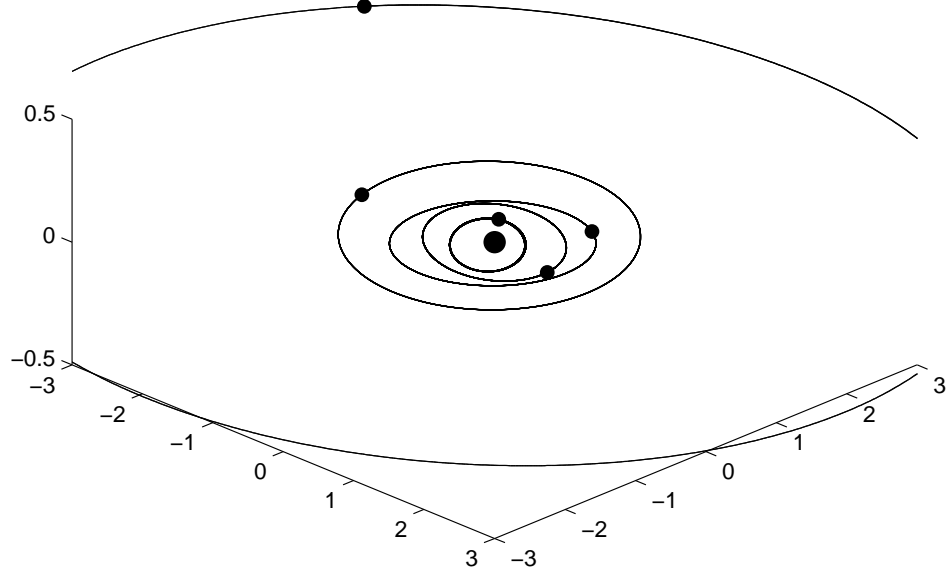
FIGURE 14. Orbital diagram for the inner solar produced by the spectral variational integrator using all 8 planets, the sun, and Pluto with 100 time steps with $h = 100$ days.
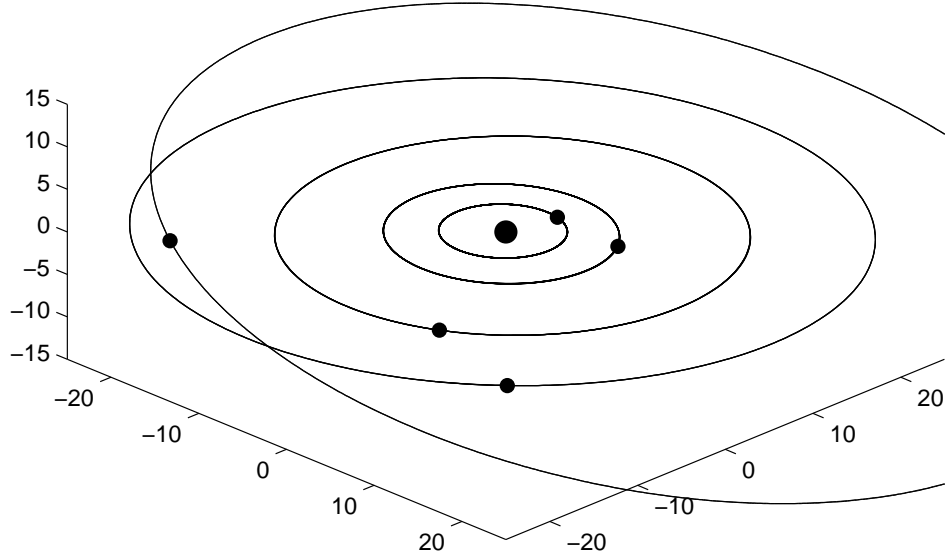

FIGURE 15. Orbital diagram for the inner solar produced by the spectral variational integrator using the 4 outer planets, Pluto, with the sun and 4 inner planets aggregated to a point with 100 time steps at $h = 1825$ days.

recovered over each time step. The error of these continuous approximations was shown to be related to the error of the one step map. Furthermore, the Noether quantities along this continuous approximation approximate the true Noether quantity up to a small error which does not grow with the number of steps taken. It was also shown that the error of the Noether quantities converges to zero with n or h refinement at a predictable rate.

In addition to the convergence results, another interesting feature of spectral variational integrators is the construction of very high order methods that remain stable and accurate using time steps that are orders of magnitude larger than can be tolerated by traditional integrators. The trade off is that the computational effort required to compute each time step is also orders of magnitude larger than that of other methods, which are a major trade off in terms of the practicality of spectral variational integrators. However, a mitigating factor of this trade off is that the approach of solving a short sequence of large problems, as opposed to a large sequence of small problems, lends itself much better to parallelization and computational acceleration. The literature on methods for acceleration of the construction and solution of structured systems of linear and nonlinear problems for PDE problems is extensive, and it is likely that such methods could be applied to spectral variational integrators to greatly improve their computation cost.

5.1. **Future Work.** Future directions for this work are numerous. Because of generality of the construction of Galerkin variational integrators, there exists many possible directions of further exploration.

5.1.1. *Lie Group Spectral Variational Integrators.* Following the approach of Leok and Shingel [16] or Bou-Rabee and Marsden [1], it is relatively straight forward to extend spectral variational integrators to Lie groups using natural charts. A systematic investigation of the resulting Lie group methods, including convergence and near conservation of Noether quantities, would be a natural extension of the work done here.

5.1.2. *Novel Variational Integrators.* The power of the Galerkin variational framework is its high flexibility in the choice of approximation spaces and quadrature rules used to construct numerical methods. This flexibility allows for the construction of novel methods specifically tailored to certain applications. One immediate example is the use of periodic functions to construct methods for detecting choreographies in Kepler problems, which would be closely related to methods already used with great success to detect choreographies. Another interesting application would be the use of high order polynomials to develop integrators for high-order Lie group problems, such as the construction of Riemannian splines, which has a variety of applications in motion planning. Enriching traditional polynomial approximation spaces with highly oscillatory functions could be used to develop methods for problems with dynamics evolving on radically different time scales, which are also very challenging for traditional numerical methods.

5.1.3. *Multisymplectic Variational Integrators.* Multisymplectic geometry (see Marsden et al. [20]) has become an increasingly popular framework for extending much of the geometric theory from classical Lagrangian mechanics to Lagrangian PDEs. The foundations for a discrete theory have been laid, and there have been significant results achieved in geometric techniques for structured problems such as elasticity, fluid mechanics, non-linear wave equations, and computational electromagnetism. However, there is still significant work to be done in the areas of construction of numerical methods, analysis of discrete geometric structure, and especially error analysis. Galerkin type methods have become a standard method in classical numerical PDE methods, such as Finite-Element Methods, Spectral, and Pseudospectral methods. The variational Galerkin framework could provide a natural framework for extending these classical methods to structure preserving geometric methods for PDEs, and the analysis of such methods will rely on the notion of the boundary Lagrangian (see Vankerschaver et al. [26]), which is the PDE analogue of the exact discrete Lagrangian.

References

[1] N. Bou-Rabee and J. E. Marsden. Hamilton-Pontryagin integrators on Lie groups. I. Introduction and structure-preserving properties. *Found. Comput. Math.*, 9(2):197–219, 2009.

[2] N. Bou-Rabee and H. Owhadi. Stochastic variational integrators. *IMA J. Numer. Anal.*, 29(2):421–443, 2009.

[3] N. Bou-Rabee and H. Owhadi. Long-run accuracy of variational integrators in the stochastic context. *SIAM J. Numer. Anal.*, 48(1):278–297, 2010.

[4] J. P. Boyd. *Chebyshev and Fourier spectral methods*. Dover Publications Inc., Mineola, NY, second edition, 2001.

[5] J. Cortés and S. Martínez. Non-holonomic integrators. *Nonlinearity*, 14(5):1365–1392, 2001.

[6] G. Dal Maso. *An introduction to Γ-convergence*. Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston Inc., Boston, MA, 1993.

[7] Y. N. Fedorov and D. V. Zenkov. Discrete nonholonomic LL systems on Lie groups. *Nonlinearity*, 18 (5):2211–2241, 2005.

[8] R. Fetecau, J. E. Marsden, M. Ortiz, and M. West. Nonsmooth Lagrangian mechanics and variational collision integrators. *SIAM Journal on Applied Dynamical Systems*, 2(3):381–416, 2003.

[9] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006.

[10] S. Lall and M. West. Discrete variational Hamiltonian mechanics. *J. Phys. A*, 39(19):5509–5519, 2006.

[11] S. Larsson and V. Thomée. *Partial differential equations with numerical methods*, volume 45 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, 2003.

[12] T. Lee, M. Leok, and N. H. McClamroch. Lie group variational integrators for the full body problem. *Comput. Methods Appl. Mech. Engrg.*, 196(29-30):2907–2924, 2007.

[13] T. Lee, M. Leok, and N. H. McClamroch. Lagrangian mechanics and variational integrators on two-spheres. *Internat. J. Numer. Methods Engrg.*, 79(9):1147–1174, 2009.

[14] M. Leok. Generalized Galerkin variational integrators: Lie group, multiscale, and pseudospectral methods. (preprint, `arXiv:math.NA/0508360`), 2004.

[15] M. Leok and T. Shingel. Prolongation–collocation variational integrators. *IMA J. Numer. Anal.*, 32(3): 1194–1216, 2012.

[16] M. Leok and T. Shingel. General techniques for constructing variational integrators. *Frontiers of Mathematics in China*, 7(2):273–303, 2012. (Special issue on computational mathematics, invited paper).

[17] M. Leok and J. Zhang. Discrete Hamiltonian variational integrators. *IMA J. Numer. Anal.*, 31(4): 1497–1532, 2011.

[18] A. Lew, J. E. Marsden, M. Ortiz, and M. West. Asynchronous variational integrators. *Arch. Ration. Mech. Anal.*, 167(2):85–146, 2003.

[19] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numer.*, 10:357–514, 2001.

[20] J. E. Marsden, G. W. Patrick, and S. Shkoller. Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Comm. Math. Phys.*, 199(2):351–395, 1998.

[21] R. McLachlan and M. Perlmutter. Integrators for nonholonomic mechanical systems. *J. Nonlinear Sci.*, 16(4):283–328, 2006.

[22] S. Müller and M. Ortiz. On the Γ-convergence of discrete dynamics and variational integrators. *J. Nonlinear Sci.*, 14(3):279–296, 2004.

[23] NASA. JPL Horizons system online solar system data and ephemeris, September 2012. URL `http://ssd.jpl.nasa.gov/?horizons`.

[24] G. W. Patrick and C. Cuell. Error analysis of variational integrators of unconstrained Lagrangian systems. *Numer. Math.*, 113(2):243–264, 2009.

[25] L. N. Trefethen. *Spectral methods in MATLAB*, volume 10 of *Software, Environments, and Tools*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

[26] J. Vankerschaver, C. Liao, and M. Leok. Generating functionals and Lagrangian PDEs. *Journal of Nonlinear Science*, 2011. (submitted, `arXiv:1111.0280 [math-ph]`.

APPENDIX A. PROOFS OF GEOMETRIC CONVERGENCE OF SPECTRAL VARIATIONAL INTEGRATORS

As stated in §3.2, it can be shown that spectral variational integrators converge geometrically to the true flow associated with a Lagrangian under the appropriate conditions. The proof is similar to that of order optimality, and is offered below.

However, before we offer a proof of the theorem, we must establish a result that extends Theorem 1.1. Specifically, we must show:

Theorem A.1. (Extension of Theorem 1.1 to Geometric Convergence) *Given a regular Lagrangian L and corresponding Hamiltonian H, the following are equivalent for a discrete Lagrangian $L_d(q_0, q_1, n)$:*

(1) *there exist a positive constant K, where $K < 1$, such that the discrete Hamiltonian map for $L_d(q_0, q_h, n)$ has error $\mathcal{O}(K^n)$,*
(2) *there exists a positive constant K, where $K < 1$, such that the discrete Legendre transforms of $L_d(q_0, q_h, n)$ have error $\mathcal{O}(K^n)$,*
(3) *there exists a positive constant K, where $K < 1$, such that $L_d(q_0, q_h, n)$ approximates the exact discrete Lagrangian $L_d^E(q_0, q_h, h)$ with error $\mathcal{O}(K^n)$.*

The proof of this theorem is a simple modification of the proof of Theorem 1.1, and is included here for completeness. For details, the interested reader is referred to [19].

Proof. Since we are assuming that the time step h is being held constant, we will suppress it as an argument to the exact discrete Lagrangian, writing $L_d^E(q_0, q_h)$ for $L_d^E(q_0, q_h, h)$. First, we will assume that $L_d(q_0, q_h, n)$ approximates $L_d(q_0, q_h)$ with error $\mathcal{O}(K^n)$ and show this implies the discrete Legendre transforms have error $\mathcal{O}(K^n)$. By assumption, if $L_d(q_0, q_h, n)$ has error $\mathcal{O}(K^n)$, there exists a function which is smooth in its first two arguments $e_v : Q \times Q \times \mathbb{N} \to \mathbb{R}$ such that:

$$L_d(q_0, q_h, n) = L_d^E(q_0, q_h) + K^n e_v(q_0, q_h, n),$$

with $|e_v(q_0, q_h, n)| \leq C_v$ on U_v. Taking derivatives with respect to the first argument yields:

$$\mathbb{F}^- L_d^n(q_0, q_h) = \mathbb{F}^- L_d^E(q_0, q_h) + K^n D_1 e_v(q_0, q_h, n),$$

and with respect to the second yields:

$$\mathbb{F}^+ L_d^n(q_0, q_h) = \mathbb{F}^+ L_d^E(q_0, q_h) + K^n D_2 e_v(q_0, q_h, n).$$

Since e_v is smooth and bounded over the closed set U, so are $D_1 e_v$ and $D_2 e_v$, yielding that the discrete Legendre transforms have error $\mathcal{O}(K^n)$. Now, to show that if the discrete Legendre transforms have error $\mathcal{O}(K^n)$, the discrete Lagrangian has error $\mathcal{O}(K^n)$, we write:

$$e_v(q_0, q_h, n) = \frac{1}{K^n}\left[L_d(q_0, q_h, n) - L_d^E(q_0, q_h)\right],$$

$$D_1 e_v(q_0, q_h, n) = \frac{1}{K^n}\left[\mathbb{F}^- L_d(q_0, q_h, n) - \mathbb{F}^- L_d^E(q_0, q_h)\right],$$

$$D_2 e_v(q_0, q_h, n) = \frac{1}{K^n}\left[\mathbb{F}^+ L_d(q_0, q_h, n) - \mathbb{F}^+ L_d^E(q_0, q_h)\right].$$

Since $D_1 e_v$ and $D_2 e_v$ are smooth and bounded on a bounded set, this implies there exists a function $d(n)$ such that

$$\|e_v(q(0), q(h), n) - d(n)\| \leq C_v,$$

for some constant C_v. This shows that the discrete Lagrangian is equivalent to a discrete Lagrangian with error $\mathcal{O}(K^n)$. We note that the equivalence is a consequence of the fact that one can add a function of h or n to any discrete Lagrangian and the resulting discrete Euler-Lagrange equations and discrete Legendre Transforms are unchanged, hence the function $d(n)$.

To show the equivalence of the discrete Hamiltonian map having error $\mathcal{O}(K^n)$ and the discrete Legendre transforms having error $\mathcal{O}(K^n)$, we recall expressions for the discrete Hamiltonian map for the discrete Lagrangian L_d and exact discrete Lagrangian L_d^E:

$$F_{L_d} = \mathbb{F}^+ L_d \circ \left(\mathbb{F}^- L_d\right)^{-1},$$

34

$$F_{L_d^E} = \mathbb{F}^+ L_d^E \circ \left(\mathbb{F}^- L_d^E \right)^{-1}.$$

Now, we make use of the following consequence of the implicit function theorem: If we have smooth functions g_1, g_2 and the sequences of functions $\{f_{1_n}\}_{n=1}^\infty$, $\{f_{2_n}\}_{n=1}^\infty$, $\{e_{1_n}\}_{n=1}^\infty$ and $\{e_{2_n}\}_{n=1}^\infty$ such that

$$f_{1_n}(x) = g_1(x) + K^n e_{1_n}(x),$$
$$f_{2_n}(x) = g_2(x) + K^n e_{2_n}(x),$$

where $\sup \{\|e_{1_n}\|\}_{n=1}^\infty < C_1$ and $\sup \{\|e_{2_n}\|\}_{n=1}^\infty < C_2$ on compact sets, then

(26) $$f_{2_n}(f_{1_n}(x)) = g_2(g_1(x)) + K^n e_{12_n}(x)$$

(27) $$f_{1_n}^{-1}(y) = g_1^{-1}(y) + K^n \bar{e}_{1_n}(y)$$

for some sequences of functions $\{e_{12_n}\}_{n=1}^\infty$, $\{\bar{e}_{1_n}\}_{n=1}^\infty$ where $\sup \{\|e_{12_n}\|\}_{n=1}^\infty < C_1$ and $\sup \{\|\bar{e}_{1_n}\|\}_{n=1}^\infty < C_2$ on compact sets.

It follows from (26) and (27) that if the discrete Legendre transforms have error $\mathcal{O}(K^n)$, the discrete Hamiltonian map does as well. Finally, if we have a discrete Hamiltonian map has error $\mathcal{O}(K^n)$, we use the identity

$$\left(\mathbb{F}^- L_d \right)^{-1}(q_0, p_0) = (q_0, \pi_Q \circ F_{L_d}(q_0, p_0))$$

where π_Q is the projection map $\pi_Q : (q, p) \to q$ and (27) to see that $\mathbb{F}^- L_d$ is has error $\mathcal{O}(K^n)$, and the identity:

$$\mathbb{F}^+ L_d = F_{L_d} \circ \mathbb{F}^- L_d,$$

along with (26) to establish that $F^+ L_d$ also has error $\mathcal{O}(K^n)$, which completes the proof. $\square$

This simple extension is a critical tool for establishing the geometric convergence of spectral variational integrators, and leads to the following theorem concerning the accuracy of spectral variational integrators.

Theorem A.2. (Geometric Convergence of Spectral Variational Integrators) *Given an interval $[0, h]$ and a Lagrangian $L : TQ \to \mathbb{R}$, let $\bar{q}$ be the exact solution to the Euler-Lagrange equations, and $\tilde{q}_n$ be the stationary point of the spectral variational discrete action*

$$L_d^S(q_0, q_h, n) = \operatorname*{ext}_{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_h}} \mathbb{S}_d\left(\{q_i\}_{i=1}^n\right) = \operatorname*{ext}_{\substack{q_n \in \mathbb{M}^n([0,h],Q) \\ q_n(0)=q_0, q_n(h)=q_h}} h \sum_{j=0}^{m_n} b_{n_j} L\left(q_n\left(c_{n_j} h\right), \dot{q}_n\left(c_{n_j} h\right)\right).$$

If:

(1) *there exists constants C_A, K_A, $K_A < 1$, independent of n, such that, for each n, there exists a curve $\hat{q}_n \in \mathbb{M}^n([0, h], Q)$ such that,*

$$\left\| (\bar{q}, \dot{\bar{q}}) - \left(\hat{q}_n, \dot{\hat{q}}_n \right) \right\| \leq C_A K_A^n,$$

(2) *there exists a closed and bounded neighborhood $U \subset TQ$ such that $(\bar{q}(t), \dot{\bar{q}}(t)) \in U$ and $\left(\hat{q}_n(t), \dot{\hat{q}}_n(t) \right) \in U$ for all t and n, and all partial derivatives of L are continuous on U,*

(3) *for the sequence of quadrature rules $\mathcal{G}_n(f) = h \sum_{j=1}^{m_n} b_{n_j} f\left(c_{n_j} h\right) \approx \int_0^h f(t) dt$ there exists constants C_g, K_g, $K_g < 1$, independent of n such that:*

$$\left| \int_0^h L(q_n(t), \dot{q}_n(t)) dt - h \sum_{j=1}^{m_n} b_{n_j} L\left(q_n\left(c_{n_j} h\right), \dot{q}_n\left(c_{n_j} h\right)\right) \right| \leq C_g K_g^n,$$

for any $q_n \in \mathbb{M}^n([0, h], Q)$,

(4) *and the stationary points $\bar{q}$, $\tilde{q}_n$ minimize their respective actions,*

then

(28) $$\left| L_d^E(q_0, q_1) - L_d^S(q_0, q_1, n) \right| \leq C_s K_s^n$$

for some constants C_s, K_s, $K_s < 1$, independent of n, and hence the discrete Hamiltonian flow map has error $\mathcal{O}(K_s^n)$.

Proof. As before, we rewrite both the exact discrete Lagrangian and the spectral discrete Lagrangian:

$$\left| L_d^E (q_0, q_1) - L_d^S (q_0, q_1, n) \right| = \left| \int_0^h L\left(\bar{q}(t), \dot{\bar{q}}(t) \right) \mathrm{d}t - \mathcal{G}_n \left(L\left(\tilde{q}_n(t), \dot{\tilde{q}}_n(t) \right) \right) \right|$$

$$= \left| \int_0^h L\left(\bar{q}(t), \dot{\bar{q}}(t) \right) \mathrm{d}t - h \sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n(c_{n_j} h), \dot{\tilde{q}}_n(c_{n_j} h) \right) \right|$$

$$= \left| \int_0^h L\left(\bar{q}, \dot{\bar{q}} \right) \mathrm{d}t - h \sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n, \dot{\tilde{q}}_n \right) \right|,$$

with suppression of the t argument. We introduce the action evaluated on the curve $\hat{q}_n$:

(29)
$$\left| \int_0^h L\left(\bar{q}, \dot{\bar{q}} \right) \mathrm{d}t - h \sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n, \dot{\tilde{q}}_n \right) \right| = \left| \int_0^h L\left(\bar{q}, \dot{\bar{q}} \right) \mathrm{d}t - \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n \right) \mathrm{d}t \right.$$

$$\left. + \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n \right) \mathrm{d}t - h \sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n, \dot{\tilde{q}}_n \right) \right|$$

(30)
$$\leq \left| \int_0^h L\left(\bar{q}, \dot{\bar{q}} \right) \mathrm{d}t - \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n \right) \mathrm{d}t \right|$$

(31)
$$+ \left| \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n \right) \mathrm{d}t - h \sum_{j=1}^{m} b_{n_j} L\left(\tilde{q}_n, \dot{\tilde{q}}_n \right) \right|.$$

Considering the first term in (31):

$$\left| \int_0^h L\left(\bar{q}, \dot{\bar{q}} \right) \mathrm{d}t - \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n \right) \mathrm{d}t \right| = \left| \int_0^h L\left(\bar{q}, \dot{\bar{q}} \right) - L\left(\hat{q}_n, \dot{\hat{q}}_n \right) \mathrm{d}t \right|$$

$$\leq \int_0^h \left| L\left(\bar{q}, \dot{\bar{q}} \right) - L\left(\hat{q}_n, \dot{\hat{q}}_n \right) \right| \mathrm{d}t.$$

By assumption, all partials of L are continuous on U, and since U is closed and bounded, this implies L is Lipschitz on U, so let L_α denote the Lipschitz constant. Since, again by assumption, $(\bar{q}, \dot{\bar{q}}) \in U$ and $\left(\hat{q}_n, \dot{\hat{q}}_n \right) \in U$, we can obtain:

$$\int_0^h \left| L\left(\bar{q}, \dot{\bar{q}} \right) - L\left(\hat{q}_n, \dot{\hat{q}}_n \right) \right| \mathrm{d}t \leq \int_0^h L_\alpha \left| (\bar{q}, \dot{\bar{q}}) - \left(\hat{q}_n, \dot{\hat{q}}_n \right) \right| \mathrm{d}t$$

$$\leq \int_0^h L_\alpha C_A K_A^n \mathrm{d}t$$

$$= h L_\alpha C_A K_A^n.$$

Hence,

(32)
$$\left| \int_0^h L\left(\bar{q}, \dot{\bar{q}} \right) \mathrm{d}t - \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n \right) \mathrm{d}t \right| \leq h L_\alpha C_A K_A^n.$$

Next, considering the second term in (31),

$$\left| \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n \right) \mathrm{d}t - \sum_{j=1}^{m} h b_{n_j} L\left(\tilde{q}_n, \dot{\tilde{q}}_n \right) \right|,$$

36

since $\tilde{q}_n$ minimizes its action,

$$(33) \qquad h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \leq h\sum_{j=1}^{m_n} b_{n_j} L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \leq \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t + C_g K_g^n$$

where the inequalities follow from the assumptions on the order of the quadrature rule and (32). Furthermore,

$$h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \geq \int_0^h L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \mathrm{d}t - C_g K_g^n \geq \int_0^h L\left(\bar{q}, \dot{\bar{q}}\right) \mathrm{d}t - C_g K_g^n$$

$$(34) \qquad\qquad\qquad\qquad \geq \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t - hL_\alpha C_A K_A^n - C_g K_g^n,$$

where the inequalities follow from (32), the order of the sequence of quadrature rules, and the assumption that $\bar{q}$ minimizes its action. Putting (33) and (34) together, we can conclude:

$$(35) \qquad \left| \int_0^h L\left(\hat{q}_n, \dot{\hat{q}}_n\right) \mathrm{d}t - h\sum_{j=1}^{m_n} b_{n_j} L\left(\tilde{q}_n, \dot{\tilde{q}}_n\right) \right| \leq \left(hL_\alpha C_A + C_g\right) K_s^{-n}.$$

where $K_s = \max\left(K_A, K_g\right)$. Now, combining the bounds (32) and (35) in (31), we can conclude

$$\left| L_d^E\left(q_0, q_1\right) - L_d^S\left(q_0, q_1, n\right) \right| \leq \left(2hL_\alpha C_A + C_g\right) K_s^{-n}$$

which, combined with Theorem A.1, establishes the rate of convergence. $\qquad\square$

Department of Mathematics, University of California, San Diego, 9500 Gilman Drive #0112, La Jolla, California 92093-0112, USA

E-mail address: j9hall@math.ucsd.edu

Department of Mathematics, University of California, San Diego, 9500 Gilman Drive #0112, La Jolla, California 92093-0112, USA

E-mail address: mleok@math.ucsd.edu